DOCUMENT RESUME

ED 056 414 24 EA 003 901

AUTHOR Welch, Wayne; And Others
TITLE Conceptual Papers Defining the Knowledge and Skills Required To Function as Educational Developers and Evaluators. Design Document II for the Midwest Educational Training Center.
INSTITUTION Upper Midwest Regional Educational Lab., Inc., Minneapolis, Minn.
SPONS AGENCY National Center for Educational Research and Development (DHEW/CE), Washington, D.C.
BUREAU NO BR-6-2870
PUB DATE 14 Dec 70
CONTRACT OEC-3-7-062870-3069
NOTE 144p.

EDRS PRICE MF-$0.65 HC-$6.58
DESCRIPTORS *Administrative Personnel; *Behavioral Objectives; Educational Development; Educational Objectives; Evaluation; Instructional Materials; Job Skills; Knowledge Level; Management Development; *Project Applications; *Regional Laboratories; Research and Development Centers; Skill Analysis; *Skill Development

ABSTRACT

This document, one part of a project to train personnel for educational development and evaluation, presents the rationale underlying the training procedure. The presentation is organized around conceptualizations of the knowledge and skills required by development and evaluation personnel. Papers treat conceptually many instructional objectives and terminal behaviors as guidelines for developers and evaluators. The objectives outlined will eventually be organized into training materials for evaluation personnel and into sequences for research, development, and diffusion. Related documents are EA 003 900, EA 003 902, and EA 003 903. (RA)

# DESIG

for the MIDWES

Upper Midwes
University of
Educational R
of the
Control Data
Southwest Sta
Iowa State De
Minnesota Sta
North Dakota
South Dakota
Wyoming Stat

APOL

DE-NCERD
EA
BR-[redacted]
G2870

# SION DOCUMENT II

VEST EDUCATIONAL TRAINING CENTER

Submitted by:

west Regional Educational Laboratory
of Minnesota
Research and Development Council
the Twin Cities Metropolitan Area
a Corporation
State College Marshall, Minnesota
Department of Education
State Department of Education
State Department of Education
ota State Department of Education
State Department of Education

Design Document II

for the MIDWEST EDUCATIONAL TRAINING CENTER

# Conceptual Papers Defining the Knowledge and Skills Required to Function as Educational Developers and Evaluators

Prepared by:

Wayne Welch
Vernon Hendrix
Paul Johnson
James Terwilliger



Upper Midwest Regional Educational Laboratory
1640 East 78th Street
Minneapolis, Minnesota 55423

December 14, 1970

Conceptual Papers Defining the Knowledge
and Skills Required to Function as
Educational Developers and Evaluators

Table of Contents

## Part 1 - OVERVIEW

The guiding principle of the proposed Midwest Educational Training Center is that training will be provided in terms of desired terminal behaviors rather than in terms of courses taken or time spent. Prerequisite to the implementation of this philosophy is the need to define precisely the desired instructional objectives for research, development, diffusion, and evaluation (R,D,D, and E) personnel. Once this task has been accomplished, training materials and sequences (called training modules in this proposal) can be developed to achieve the specified objectives. Implementation of the training packages on an individual basis should result in the desired terminal behaviors. Achievement of the set of terminal behaviors considered essential for the roles identified in Design Document I for each trainee is the ultimate purpose of the consortium.

In Design Document II we have specified a large number of instructional objectives and terminal behaviors for educational developers and evaluators. The training modules in the form of conceptual papers will draw partially upon the objectives stated in this document. Because these papers were written by experienced researchers, developers, and evaluators, there are slight variations in style and focus. However, the ultimate purpose of the documents, to define the knowledge and skills for the trainees, has been accomplished. From this wealth of objectives will be drawn the components of the training modules.

A word of explanation is in order for the decision to focus on Development and Evaluation, and to largely assign Diffusion and Research lower priority in our consortium. The decision was reached on the basis of our survey of manpower needs in the R,D,D, and E areas, our belief that research training is and has been in general adequately supplied, and some doubt in our own minds about the capabilities of our consortium to provide diffusion training experience. This coupled with a belief, so aptly confirmed in the Worthen and Glass AERA paper (1970), resulted in the decision not to attempt to train diffusion personnel for the time being. Instead we will use our individualized instructional approach in the more important areas (as we perceive them) of development and evaluation.

Accordingly, the component parts of Design Document II are conceptualizations of the knowledges and skills required by development and evaluation personnel. The set of skills defined therein are coded and will comprise the set of behaviors to be accomplished by the training modules. It should be pointed out that our lists are not exhaustive and will be further enriched during the implementation stage. We hope to draw on the relevant work being conducted by the Schalock study and the AERA task force. Also the completed conceptual papers by Briggs, Glass and Worthen, Gideonse, and Schalock and Sells should be valuable inputs to our objectives pool.

One final word in this overview. Our initial efforts included research training as part of the proposal. A conceptual paper describing the objectives of such training was written by a member

of our staff, Dr. Paul Johnson. With the decision not to focus on Research training, his admirable efforts were somewhat diminished. However, we believe the tone and content of his paper will still be useful, particularly in developing the evaluation training modules, and are including it as an Appendix of Design Document II. Certainly in the overlap areas between research and evaluation, his set of objectives will be useful during the implementation stage.

The three conceptual papers that follow focus on Development, Evaluation, and Research respectively. Lists of training objectives that comprise the set of training behaviors follow each of the conceptual papers.

## Part 2 - OBJECTIVES FOR EDUCATIONAL DEVELOPMENT

### A. Introduction and Rationale

In the early stages of developing our proposal, the staff decided that it should be concerned not only with the development of a proposed "instructional shop" but simultaneously would be training people to occupy roles in instructional shops in general. The term "instructional shop" is considered quite broadly here and could include currently existing and quite familiar educational institutions such as elementary schools, secondary schools, college departments, etc. This term also includes components of more complex organizations, such as industry and the military, which are concerned primarily with training of personnel. Many examples of such instructional shops might be cited among larger industries such as International Business Machines, Control Data Corporation, American Telephone and Telegraph, and the various training schools in the Armed Forces. This broad definition would also include private instructional concerns such as Deterline and Associates, Tiemann Associates, J. H. Harless, Inc., General Programmed Teaching, Inc., Dorset Enterprises, and similar devisions of larger corporations such as Xerox, Westinghouse, and the Rand Corporation.

One of the staff members of this planning project was personally familiar with three of the above mentioned "shops," and to gain further knowledge made a visit to the headquarters of the Field Engineering Instructional System of the IBM Corporation in Poughkeepsie, New York. Conversations with personnel at IBM and examination of documents provided by them have contributed substantially to the creation of this document.

A brief description of the operation of the IBM Field Instructional System might be useful as a background for the "shop" concept and the roles and objectives which will be listed. The Field Instructional System is concerned with rather specific tasks originated by the introduction of new hardware, operating systems, changes in hardware or systems, etc. For example, as a new product is being developed by the corporation, a course development representative (or several depending on the complexity of the product) would be assigned to work with the scientists and engineers developing the project. If the product should be something rather complicated such as the development of a new computer system, this "assignment" might occupy several years. It is the purpose of the course development representatives to specify the tasks, skills, and concepts which personnel servicing the product must master. The objectives defined by this person would correspond approximately to the instructional objectives identified in the present document. The course development representative, who must possess some "subject matter" knowledge himself relating to the product, must "satisfy" the subject matter experts (scientists and engineers) who are developing the product, that a person with the competencies he has specified can, in fact, adequately service the product. The work of this person is then turned over to other people in the Field Instructional System who develop the specific behavioral objectives and consider ways in which instructional modules might be produced to allow field engineers to acquire these behaviors. After their work is completed a feedback loop to the scientists and engineers is introduced to again satisfy the system that the appropriate behavioral objectives

have been specified. Also, decisions are made at this point as to the exact form of the instruction (classroom, lecture, textbook, laboratory, programmed materials, or computer assisted instruction). The primary factors considered at this management decision point are available resources and costs of providing the instruction[1] (including development of CAI programs or, for example, bringing small groups of personnel to central locations for extended instruction). At this checkpoint the experts are also provided with the criterion measures by which the instruction will be evaluated and upon which the certification of the learners will be based. This must also satisfy them. Next, the results of this step in the process are then turned over to as many "authors" as needed, to develop the actual instructional modules. After these are developed and thoroughly tested with samples of learners, feedback is again provided to all of the earlier stages in the system, including the subject matter experts, to verify that the correct "products of instruction" are being achieved and this will result in adequate performance on the job. At this point the actual instructional module is initiated. As the training procedes, further "bugs" are identified and worked out of the module, again with feedback loops to earlier steps in the process.

This example, although it relates to some of the sections in this document, should not be regarded as limiting or the only type of system for which this proposed training program might provide

---

1. Reference to the management scheme found elsewhere in the proposal will show that many of these same decisions have been included, although the proposed management scheme was developed independently of the study of the IBM system.

personnel. The management system provides for complete flexibility in the extent to which trainee needs, employer requirements, and limitations of the instructional system can be interfaced and analyzed. We propose to be able to train people with highly differentiated and specialized skills that might function in a shop similar to that described above. This "shop" could be in an industrial organization or in an educational institution. Also, our management system would provide for a "one man shop" if needed, and could produce a person with a variety of skills in all of the designated areas. Recent literature including Showmaker[2] and Deterline[3], as well as instructional practice in business and industry (which some might argue is more effective and progressive than that found in many public educational institutions) indicates that the greater need is likely to be for training personnel with differing developmental skills. Accordingly, based on a synthesis of suggestions in the literature, and a study of instructional shops, four types of developmental roles are defined. These roles are: Curriculum Development Specialist, Instructional Development Specialist, Educational Systems Management Specialist, and Evaluation Specialist.

It is recognized that in any operational instructional system there must necessarily be a high degree of coordination across the four role areas that have been defined. This is especially

---

2. Harry A. Shoemaker, "Instructional Technologists: Who Needs Them?," National Society for Programmed Instruction Journal, vol. 8, no. 5, May, 1969.

3. William A. Deterline, "Practical Problems in Program Production," in Programmed Instruction, 66th Yearbook of the N.S.S.E., Part II, Phil C. Lange (ed.), Chicago: U. of Chicago Press, 1967, pp. 178-216.

true of Curriculum Development and Evaluation. As objectives are developed for a new course or program, simultaneously a consideration must be given to methods of determining the achievement of these objectives and placing values upon the objectives themselves. Also, since instructional development done correctly requires a great expenditure of time, resources and money, it must be directly coordinated with evaluation and management (in terms of determining values and costs). An instructional developer should not be assigned the development of significant instructional materials until, by the best methods possible, it has been determined that (1) the content analysis is correct, (2) the intended student has use for such content, (3) devoting time to the content does not interfere with more valuable learning, and (4) the amount of time to be allotted appears to permit adequate attainment of the objectives by the student.

It must then be recognized that in any operational instructional system, all of the skills must either be present in one person, or a team of persons with differentiated skills must be closely coordinated. This again underscores the importance of the management role.

Another type of coordination must be assumed. This refers to the necessary coordination between curriculum development, instructional development, evaluation, and subject matter experts. Again, either the curriculum and instructional developers and evaluators must themselves possess the necessary subject matter expertise, or they must work closely in coordination with a subject matter expert. In the case where they must work with the subject matter expert, they must have the necessary skills to do this. It will be noticed

in the listing of objectives to follow, that references are made to such skills.

With this rationale, we will now move to enlarge upon three of the four roles specified in the developmental process. These are designated as Curriculum Development Specialists, Educational Systems Management Specialists and Instruction Development Specialists. E ational Evaluation Specialists are discussed in Part 3 of this document. Within each of these role areas a number of conceptual objectives are presented. Each of these might be considered a task cluster, and would probably be most analygous to currently existing positions in educational institutions. These conceptual objectives or task clusters are generally arranged within each of the role areas in a decreasing complexity of required skills and abilities. The listing of the objectives was derived from a review of the literature and self-analysis conducted by persons currently involved in the various developmental areas.

It should be recognized that the organization of objectives by role and task clusters introduces a significant amount of overlap. For example, objectives which might be specified for Curriculum Development Specialists, which require a knowledge of learning Psychology, would be quite similar if not identical to objectives specified for Instruction Development Specialists in the same areas. Similarly, the statistical skills required for Evaluation Specialists would overlap the statistical skills of researchers in all three of the other role areas, especially Educational Systems Management Specialists. The proposed Instructional Management System would account for these duplications and produce the dependencies necessitated by the individual

input and output characteristics of learners and the behavioral objectives themselves.

The specification of all specific behavioral objectives is beyond the scope of this planning project. For example, some instructional objectives, which have quite specific behaviors, refer to concepts which might be analyzed and defined in a number of different ways, thus resulting in different behavioral objectives. In this paper, a broad set of objectives is listed. When the training modules are developed during the implementation stage, they will use the set of objectives herein reported. The exact specification of modules to objectives will occur later.

In the listings of objectives for the various roles that follow, the specification of the training objectives of the program are carried no further than the "instructional objectives" level. It is intended that the instructional objectives will correspond, in most cases, to individual training modules. These might consist of pieces of programmed instruction, parts of courses or at times complete courses, practical experiences with specific outcomes in mind for the trainee, etc. It is likely, of course, that a few of the more complicated instructional objectives might consist of several parts or sub-modules, which the trainee would proceed through in a sequential order.

B. Role by Objective Matrix

I. Curriculum Development Specialists

A. Research resulting in improved curriculum theory.

1. Identifying sources and procedures for deriving objectives.
2. Conceptualizing the "structure of knowledge" in various subject areas or disciplines.
3. Identifying the interaction of education and instructional technology, with other domains (Social and Psychological behaviors, and structures, Philosophy, etc.).

B. Research resulting in improved concept behavioralization and analysis procedures for conceptual learning.

1. Identifying techniques for operationally defining conceptual behavior.
2. Operationalizing cognitive behavior (conceptual behavior, cognitive "styles," problem solving behavior, creativity, learning, etc.).

C. Research resulting in improved procedures for specifying objectives.

1. Identifying linguistic and semantic problems of communicating concepts.
2. Identifying techniques for operationalizing effective and psychomotor behavior.
3. Identifying and improving task and skill analysis procedures.

D. Research resulting in improved procedures for validating objectives.

1. Identifying techniques for placing values on educational outcomes (societal, individual, organizational, etc.).
2. Developing techniques for recognizing and predicting instructional needs (for individual, society, and organizations).

E. Specification of objectives to be achieved by learner--produced by instruction.

> For desired educational outcomes representing any type of learning, i.e., cognitive (e.g., associations, mulitiple discriminations, behavior chains, concepts, principles, process skills -- strategies, etc.), psychomotor, or affective learning, the Instructional Development Specialists will specify a set of operationally specific (behavioral) objectives acceptable to the SME (subject matter expert). Objectives are operationally specific "when a student performance test -- perpared for the objective by a second individual or group also knowledgeable in the subject matter area -- would result in observable student behavior acceptable to the first individual or group. Acceptability consists of two factors. Both the level of proficiency indicated by the observable behavior and the conditions under which such behavior would be exhibited must be acceptable." (Tiemann, 1967).
>
> The "set" of operationally specific objectives must cause the SME to agree that acceptable performance of the set would constitute evidence that the student had attained the desired educational outcome--no matter how broad and all-inclusive that outcome might be.

1. Specifying the terminal behaviors (performances) of learners who should possess designated skills (competencies).
   a. The subject will discriminate between objectives stated in specific behavioral form and objectives which are not so stated.
   b. Given a behavioral classification scheme, the subject will classify specific behavioral instructional objectives according to the scheme.

c. Given a behavioral classification scheme, the subject will write exemplary objectives in a given subject matter or content area.

d. Given generally stated objectives, the subject will restate these in specific behavioral form.

e. Given that the learner population must perform specified tasks, the subject will perform a task analysis and specify terminal objectives for a training program.

f. Given a loosely stated educational goal, the subject will probe the stater's behavior to obtain further specificity. (The stater could be himself, but not necessarily.)

g. Given a relatively specific description of what learners are expected to do, the subject will determine whether the statement implies a behavior management problem or an instructional problem.

h. Given a behaviorally stated objective, the subject will select those parts of it requiring further analysis techniques, e.g., illustrate the range (or generative characteristics) of the types of specific objectives to be classified.

i. Given a specification of terminal performance, the subject will do a deficiency analysis, enabling him to determine the instructional problem to be solved.

j. Given a set of potential course objectives, the subject will employ various heuristics for determining the "worth" of these objectives (values).

k. Given a set of potential course objectives, the subject will employ vari ; other heuristics for determining the attainability of such objectives (feasibility).

l. Given a production task (essay writing, etc.) the subject will analyze it into its components -- the response learnings involved, the discriminations involved, attitudinal factors, cognitive processing skills, and so forth.

m. Given a production task (such as essay writing or furniture building etc.), the subject will develop an appropriate checklist for evaluating the product.

n. The subject specialists will demonstrate sufficient familiarity with behavior shaping techniques such that the SME, once he realizes what an exacting and time-consuming task he has committed himself to, will persist until the appearance of the first tangible evidence of instructional improvement (usually the student behavior after exposure to the first draft or validated instruction).

o. The subject specialists will operate in a nonaversive manner, hopefully from a data-base, in order to cause the policy-makers to question basic assumptions upon which many administrative procedures are based.

2. Analyzing conceptual learning.

a. Given a "course" or "unit" sized chunck of conceptual subject matter, the subject will identify the concepts contained in it.

b. In conjunction with a subject matter expert (or group thereof), the subject will determine the critical attributes of each concept, the irrelevant attributes of typical examples, and a set of close-in non-examples of the concept.

c. On the basis of the above analysis of each concept, the subject will construct a conceptual hierarchy relating each concept to the others in terms of subordinate and superordinate relationships.

d. Given a set of typical students, the subject will devise measures of the familiarity of points in the hierarchy, by testing recognition of examples and non-examples of each concept. (It is not the case that low-level concepts are necessarily grasped before high-level ones, despite the "logic" of such a statement).

e. Based on the results of the empirical data coordinated with the logical analysis, the subject will "logically" sequence instruction in unknown concepts by proceeding from familiar to unfamiliar. Some of these sequences will involve training of discriminating subclasses within a known larger class (e.g., classes of "insects"). Other sequences will involve training in generalizing to a broader class made up of known subclasses (e.g., "herbivores" from cows and deer).

f. Based on the logical analysis, the subject will construct an appropriate conceptual learning criterion consisting of new examples and non-examples of each

class. (Actually, they are "new" in the sence that they will now be kept out of the instructional sequence and reserved for the testing sequence).

g. In designing content for teaching frames the subject will select widely differing examples and juxtapose these to create maximum generalization.

h. In designing content for teaching frames, the subject will select the closest non-examples possible for each critical attribute of the class, and juxtapose these to create fine discriminations.

i. Depending on his knowledge of the verbal sophistication of the target students, the subject will adjust the verbalization of the differences between examples and non-examples and the similarity between examples and further examples to as simple a level as required.

j. The subject will reserve further examples and non-examples for use in an expanded program should developmental testing demonstrate that the teaching is insufficient.

k. In developmental testing, the subject will probe for student difficulties with irrelevant attributes not located by subject matter experts but proving distracting to students and with distinctions which prove too difficult.

l. The subject will check each expansion of the program against the items reserved for the test to determine that he is still holding to the basic conceptual objective and is not "teaching the test."

3. Specifying the terminal behaviors (performances) of learners who should "know" designated concepts.

   a. The subject will discriminate between objectives stated in specific behavioral form and objectives which are not so stated.

   b. Given a behavioral classification scheme, the subject will classify specific behavioral instructional objectives according to the scheme.

   c. Given a behavioral classification scheme, the subject will write exemplary objectives in a given subject matter or content area.

   d. Given generally stated objectives, the subject will restate these in specific behavioral form.

   e. Given that a learner population should "understand" a stated concept, the subject will perform a concept analysis and specify terminal behavioral objectives of a training program, the achievement of which will satisfy the subject matter expert.

   f. Given a loosely stated educational goal, the subject will probe the stater's behavior to obtain further specificity. (The stater could be himself, but not necessarily.)

   g. Given a relatively specific description of what learners are expected to do, the subject will determine whether the statement implies a behavior management problem or an instructional problem.

h. Given a behaviorally stated objective, the subject will select those parts of it requiring further analysis techniques, e.g., illustrate the range (or generative characteristics) of the types of specific objectives to be classified.

i. Given a specification of terminal performance, th subject will do a deficiency analysis, enabling nim to determine the instructional problem to be solved.

j. Given a set of potential course objectives, the subject will employ various heuristics for determining the "worth" of these objectives (value).

k. Given a set of potential course objectives, the subject will employ various other heuristics for determining the attainability of such objectives (feasibility).

l. The subject specialist will demonstrate sufficient familiarity with behavior shaping techniques such that the SME, once he realized what an exacting and time-consuming task he has committed himself to, will persist until the appearance of the first tangible evidence of instructional improvement (usually the student behavior after exposure to the first draft of validated instruction).

m. The subject specialists will operate in a nonaversive manner, hopefully from a data-base, in order to cause the policy-makers to question basic assumptions upon which many administrative procedures are based.

F. Specification of learner characteristics.

1. Determining if learners possess the minimum prerequisite competencies for a designated instructional module.

a. Given an instructional module, the subject will state a terminal objective intended by the module.

b. Given an instructional module, the subject will specify the minimum pre-learning behaviors required of learners.

c. Given instructional objectives and sample measurement procedures, the subject will discriminate between appropriate and inappropriate measurement procedures.

d. The subject will discriminate between objectives stated in specific behavioral form and objectives which are not so stated.

e. Given a behavioral classification scheme, the subject will classify specific behavioral instructional objectives according to the scheme.

f. Given a behavioral classification scheme, the subject will write exemplary objectives in a given subject or content area.

g. Given generally stated objectives, the subject will restate these in specific behavioral form.

h. Given that a learner population must perform specified tasks, the subject will perform specific analysis and specify a terminal objective among a training program.

i. Given that a learner population should "understand" a stated concept, the subject will perform a concept analysis and specify terminal behavioral objectives of

a training program, the achievement of which will satisfy subject matter experts.

j. Given a specific behavioral objective, ths subject will produce exemplary criterion measures for each of the categories in a given classification of criterion measures.

k. Given specific behavioral objectives, the subject will develop appropriate criterion measures.

l. Given a classification scheme for criterion measures and sample measures, the subject will correctly classify the sample measures.

m. Given a specific behavioral objective and content, along with sample prerequisite skills, the subject will discriminate between essential and non-essential prerequisite skills.

n. Given a specific behavioral objective and content, the subject will specify in terms of behavior and content all prerequisite skills required of learner.

o. Given a discrimination task, the subject will determine whether it is one-directional or two.

p. Given a two-way discrimination (association) task, the subject will construct appropriate criterion measures for each.

q. Given a discrimination learning problem, the subject will construct criterion measures at different levels of difficulty.

r. Given a criterion measure, the subject will analyze it into component behaviors.

s. Given a statement of prerequisites, the subject will construct a test of these, validate that test, and then determine whether such prerequisites are rational for the general population to be served. (General objective: since not all the answers are in on instructional strategies and even when they are, they could be replaced by more elegant solutions). All subjects should exhibit a strong tendency to keep in touch with new developments (not to be confused with research literature).

t. Given a set of rules or principles to be mastered, the subject will determine the kinds of components involved in the rule (concepts, identities, multiple situations) and the devise tests for student mastery of the components.

2. Specifying performances which will sample and represent the possession of designated competencies.

a. Given generally stated objectives, the subject will restate these in specific behavioral form.

b. The subject will discriminate between objectives stated in specific behavioral form and objectives which are not so stated.

c. Given a behavioral classification scheme, the subject will classify specific behavioral instructional objectives according to the scheme.

d. Given a behavioral classification scheme, the subject will write exemplary objectives in a given subject matter or content area.

e. Given a classification scheme for criterion measures and sample measures, the subject will correctly classify the sample measures.

f. Given a specific behavioral objective, the subject will produce exemplary criterion measures for each of the categories in a given classification of criterion measures.

g. Given instructional objectives and sample measurement procedures, the subject will discriminate between appropriate and inappropriate measurement procedures.

h. Given specific behavioral objectives, the subject will develop appropriate criterion measures.

i. Given a discrimination task, the subject will determine whether it is one-directional or two.

j. Given a two-way discrimination (association) task, the subject will construct appropriate criterion measures for each.

k. Given a discrimination learning problem, the subject will construct criterion measures at different levels of difficulty.

l. Given a criterion measure, the subject will analyze it into component behaviors.

II. Educational Systems Management Specialists

A. Research resulting in improved organizational analysis methods.

1. Identifying structures of organizations (control, influence, decision making, etc.).

2. Measuring communications networks in organizations.

3. Measuring reward and motivation systems in organizations.

B. Research resulting in improved systems analysis and management procedures.

1. Developing quantitative decision-making techniques.

2. Achieving congruence of rational/scientific systems structures with social/psychological structures in organizations.

3. Developing computer/man interface procedures (hardware, computers, operating systems, languages, etc.).

C. Creation of instructional systems.

1. Designing management systems to control and monitor the movement of learners into and through a set of instructional modules.

a. Given an instructional module and criterion measures, the subject will construct a decision system to monitor the progress of the learner through the module.

b. Given a set of instructional modules described for a population of learners, the subject will construct a decision system which identifies those modules individual learners require and may take (based on diagnostic testing).

c. Given the elements of a decision system with all dependencies specified, the subject will prepare PERT type analyses.

2. Determining the costs-benefits of alternative instructional modules for sets of learners with specified instructional needs.

   a. Given the specifications for an instructional module, the subject can estimate the costs of producing the module in different formats.

   b. Given that instructional outcomes have been specified for learner populations, a subject can estimate the cost (individual, resources, organizational, developmental, etc.) of applying alternative instructional modules.

   c. The subject will distinguish between the cost factors attributable to the functions of production in or distribution by any particular medium.

   d. The subject will distinguish between fixed and variable media costs with respect to either the production or the distribution function.

D. Management of instructional systems.

1. Collecting reliable data on learner achievement and diagnostic testing.

   a. Given educational prescriptions for a learner, the subject can select from within a system the correct diagnostic and achievement tests to be administered.

   b. Given the results of achievement and diagnostic testing for a learner, the subject can specify the correct instructional modules and starting points for an individual learner.

2. Collecting reliable data on criterion performance of learners at decision points in a system.

   a. Given that a learner is progressing through a set of instructional modules, the subject can identify those points when criterion measures must be administered and select the correct criterion measures.

   b. Given that a selected criterion measure must be administered, the subject can administer the test and identify the correct decisions for following instruction, based on the system.

3. Processing data and computing statistics.

   a. Given a particular test, the subject can compute needed statistics (reliability, validity, error rate, required probabilities, etc.).

   b. Given that tests for decision purposes have been administered to a learner progressing through an instructional system, the subject can correctly compute necessary statistics (Bayesian probabilities of mastery, probability of successful completion, etc.).

E. Operation of instructional systems.

1. Administering instructional modules (teachers, laboratory assistants, CAI systems operators and monitors, etc.).
2. Administering achievement, criterion, and diagnostic tests.
3. Maintaining student and financial accounting records (manually or with computers).
4. Assisting professionals (teacher aides, clerks, key punchers, machine operators, computer operators, classroom monitors, secretaries, etc.).

III. Instructional Development Specialists

A. Research resulting in improved learning/instructional theory.

1. Operationalizing the categories of behaviors called learning.

2. Differentiating the conditions which determine various learning behaviors.

B. Research resulting in improved learning/instructional technology.

1. Identifying the most effective learning strategies for specified objectives, learning behaviors and learners.

2. Increasing the predictability of manipulations of learning strategies under varied conditions.

C. Design of instructional packages (courses, modules, programs, etc.). (It is assumed that most of the educational and instructional objectives listed under Roman numeral I, capital letters E and F, would constitute many of the necessary prerequisite behaviors for this conceptual objective).

1. Specifying and sequencing intermediate objectives to take learners with specified prerequisite competencies to achievement of terminal objectives.

a. Given a set of terminal objectives, the subject can select the most effective learning strategy (chaining, etc.).

b. Given a set of instructional objectives, the subject will sequence these in the most effective and/or efficient way.

c. Given a relatively specific description of what learners are expected to do, the subject will determine whether the statement implies a behavior management problem or an instructional problem.

2. Selecting appropriate content for objectives specified for an instructional module.

a. Given specific behavioral objectives, the subject will generate rules for deciding upon appropriate and inappropriate content for the instructional module.

b. Given specific behavioral objectives, the subject will discriminate between appropriate and inappropriate content for the objectives.

c. Given a specific behavioral objective, the subject will describe examples of appropriate content.

D. Creation of instructional packages. (It is assumed that most of the educational and instructional objectives in Roman numeral I, capital letters E and F, and in capital letter C above, would constitute the necessary prerequisite behaviors for this conceptual objective).

1. Given a student of generally appropriate characterisitcs, the subject will determine if the student has the appropriate entry behaviors and lacks the terminal behaviors, being therefore a good subject for developmental testing.

a. Given a set of criterion measures, the subject will try these out on appropriate students to be sure that they can be achieved by those who "know" and can't by those who don't.

b. Given a set of objectives and enough time to do so, the subject will design the leanest possible draft of materials to achieve the objectives and let student difficulties and queries determine the expansion at appropriate points.

c. Given data on differential student achievement of prerequisites, the subject will construct an individualized tract leading up to the main track.

d. Given a set of items to be memorized by students, the subject will design an instructional system which will permit individualization of practice according to student need.

e. Given a set of objectives and an intended population of learners, the subject will determine an appropriate level of student achievement beyond which he will cease developmental testing. (Being reasonable, he may raise or lower this level based on practical experience and cost requirements during development.)

f. Given a problem-solving process to be mastered, the subject will construct exercises which require the process to be applied to already-known phenomena.

2. Sequencing an objective most effectively.

a. Given a terminal objective and intermediate objectives, the subject will sequence the intermediate objectives in the most effective manner.

b. Given a terminal objective and all necessary intermediate objectives, the subject can specify all of the dependencies among the intermediate objectives and indicate necessary branching.

c. Given a first draft of instructional materials, the subject will present this to an appropriate student in such a way as to maximize feedback about the materials from the student.

d. The subject will handle student errors and student criticism in a completely supportive manner, encouraging rather than discouraging student feedback.

e. The subject will probe student difficulties in such a way as to determine the source of the problem in the materials without giving away the appropriate response.

f. Given a tape or transcript of student feedback, the subject will make changes in the instructional materials which retain the original objectives but reduce the likelihood of error in the next student.

g. Given a set of discriminations to be learned, the subject will arrange them according to least potential for interference, of the learning ability of the student makes this possible.

h. Given a set of component behaviors, the subject will determine whether a logical hierarchy is implied in sequencing these for instruction.

i. Given a logical hierarchy, the subject will validate its existence by appropriate tryouts with students.

j. Given a set of objectives with no predetermined sequence, the subject will select an appropriate sequence if time does not permit further exploration with students.

k. If time permits further exploration, the subject will determine which sequencing appeals most to students by observing their inquiry behaviors.

l. Given a principle to be mastered, the subject will sequence it in the instruction in such a way that almost all elements in it have been mastered. (No data exist on how many elements might be mastered concurrently in complex principles in physics or grammar -- but we would say not many!)

3. Specifying appropriate reinforcers and consequences (contingency).

a. Given an instructional objective and a set of reinforcers, the subject will discriminate between the appropriate and inappropriate reinforcers, based on the characteristics of a population of learners.

b. Given the results of the administration of instructional modules to a group of learners, the subject will develop more effective reinforcement schedules.

c. Given an instructional objective, the subject will specify effective and/or efficient reinforcers, based on the relevant characteristics of a population of learners.

d. Given a student or group of students going through a draft of instructional materials, the subject will observe all signs of boredom and disinterest and will probe for student attitude in post-interviews.

4. Selecting appropriate media.

a. Given an instructional objective, and a set of instructional procedures using different media, the subject will discriminate between more appropriate and less appropriate uses of media.

b. The subject will select an appropriate medium, given the demands imposed by the task analysis in interaction with the available funds, the on-going system, and the motivational requirements.

c. Given the cheapest medium which the instruction might operate (e.g., straight print, for instance), the subject will prepare alternate suggestions for increasing the motivational impact.

d. The subject will make first-draft adaptations of timing, vocabulary and sentence length, and question difficulty appropriate to information presentation in an oral medium.

e. The subject will select instructional media which optimize the benefit of all resources, i.e., cost of media, cost of instructional staff, and cost of student time committed to the instructional effort.

f. The subject will identify media by type (and combination of types) of presentation capabilities (stimulus control), i.e., discriminations or generalizations on the basis of

three-dimensional properties, sound, motion, tactile, color, odor, taste, or verbal description. As used here, 3-D refers to all aspects of spatial and relative location. Tactile includes texture, weight or mass, and psychomotor stimuli.

g. Given an operationally specific objective, the subject will list the available instructional media capable of presenting the types of stimulus (or combination of stimuli) to which the student must attend, as specified by the indicator behavior and conditions of the specific objective.

h. Having listed available media for a given specific objective, the subject will consider all factors and select for first draft trial the medium (or combination of media) which, on a best judgment basis, has a reasonable probability of conveying the instructional intent.

i. During developmental testing of instruction, the subject will employ a "lean programming" rationale with respect to mediated instruction, i.e., utilize prompting and other techniques appropriate to the media selected in order to direct attending-to behaviors to the salient aspect according to instructional intent.

j. The subject will recognize the point at which further efforts to "prompt" the initial selection of media are uneconomical during developmental testing and, at that point, shift to the "mere representative" listed media along the cost continuum.

k. The subject will evaluate the level of "media sophistication" of the SME, basing initial media selection on this level -- as one factor of cost -- in view of the time demand required as the SME interacts with media production personnel.

l. The subject will control the understandable tendencies of graphics personnel to create "mini-works-of-art" when the limits of the instructional situation, i.e., first draft trial or small number of students, do not justify an unreasonable expenditure of talent and production resources.

5. Specifying effective response modes

a. Given a specific behavioral objective and content, the subject will construct elicitors, cues, problems, etc.

b. Given a specific behavioral objective and content, the subject will describe appropriate responses.

c. Given the essential specifications for instructional modules (objectives, content, criterion measures, elicitors, cues, problems, etc), a subject will write effective instructional frames.

d. Given summary data from the administration of an instructional module, the subject will identify ineffective frames, sequencing, stimuli, etc.

e. Given a first draft of instructional materials, the subject will present this to an appropriate student in such a way as to maximize feedback about the materials from the student.

f. The subject will handle student errors and student criticism in a completely supportive manner, encouraging rather than discouraging student feedback.

g. Given a student or group of students going through a draft on instructional materials, the subject will observe all signs of boredom and disinterest and will probe for student attitude in post-interviews.

6. Selecting appropriate stimuli.

a. Given an operationally specific objective, the subject will list the available instructional media capable of presenting the type of stimulus (or combination of stimuli) to which the student must attend, as specified by the indicator behavior and conditions of the specific objective.

C. Summary List of Development Objectives

Because of overlap, the preceding list of objectives for each developmental role is longer than is essential. Therefore, a listing of each objective has been written and is attached. The objectives are numbered for easy referencing. For convenience, the objectives also have been keyed to their original developmental roles.

1. The subject will discriminate between objectives stated in specific behavioral form and objectives which are not so stated.

   Role Location: I E1a
   I E3a
   I F1d
   I F2b

2. Given generally stated objectives, the subject will restate these in specific behavioral form.

   Role Location: I E1d
   I E3d
   I F1g
   I F2g

3. Given a behavioral classification scheme, the subject will classify specific behavioral instructional objectives according to the scheme.

   Role Location: I E1b
   I E3b
   I F1e
   I F2c

4. Given a behavioral classification scheme, the subject will write exemplary objectives in a given subject matter or content area.

   Role Location: I E1c
   I E3c
   I F1f
   I F2d

5. Given that the learner population must perform specified tasks, the subject will be able to perform a task analysis and specify terminal objectives for a training program.

   Role Location: I E1e
   I F1h

6. Given a loosely stated educational goal, the subject will probe the stater's behavior to obtain further specificity. (The stater could be himself, but not necessarily.)

   Role Location: I E1f
   I E3f

7. Given a relatively specific descri ion of what learners are expected to do, the subject will rmine whether the statement implies a behavior management problem or an instructional problem.

   Role Location: I C1c
   I E1g
   I E3g

8. Given a behaviorally stated objective, the subject will select those parts of it requiring further analysis techniques, e.g., illustrate the range (or generative characteristics) of the types of specific objectives to be classified.

   Role Location: I E1h
   I E3h

9. Given a specification of terminal performance, the subject will do a deficiency analysis, enabling him to determine the instructional problem to be solved.

   Role Location: I E1i
   I E3i

10. Given a set of potential course objectives, the subject will employ various heuristics for determining the "worth" of these objectives (value).

    Role Location: I E1j
    I E3j

11. Given a set of potential course objectives, the subject will employ various other heuristics for determining the attainability of such objectives (feasibility).

    Role Location: I E1k
    I E3k

12. Given a production task (essay writing, etc.), the subject will analyze it into its components -- the response learnings involved, the discriminations involved, attitudinal factors, cognitive processing skills, and so forth.

    Role Location: I E11

13. Given a production task (such as essay writing or furniture building, etc.), the subject will develop an appropriate checklist for evaluating the product.

    Role Location: I E1m

14. The CD specialist will demonstrate sufficient familiarity with behavior shaping techniques such that the SME, once he realizes what an exacting and time-consuming task he has committed himself to, will persist until the appearance of the first tangible evidence of instructional improvement (usually the student behavior after exposure to the first draft of validated instruction).

    Role Location: I E1n
    I E31

15. The CD specialist will operate in a nonaversive manner, hopefully from a data-base, in order to cause the policy-makers to question basic assumptions upon which many administrative procedures are based.

    Role Location: I E1o
    I E3m

16. Given a "course" or "unit" sized chunk of conceptual subject matter, the subject will identify the concepts contained in it.

    Role Location: I E2a

17. In conjunction with a subject matter expert (or group thereof), the subject will determine the critical attributes of each concept, the irrelevant attributes of typical examples, and a set of close-in non-examples of the concept.

    Role Location: I E2b

18. On the basis of the above analysis (IE2c) of each concept, the subject will construct a conceptual hierarchy relating each concept to the others in terms of subordinate and superordinate relationships.

    Role Location: I E2c

19. Given a set of typical students, the subject will devise measures of the familiarity of points in the hierarchy, by testing recognition of examples and non-examples of each concept. (It is not the case that low-level concepts are necessarily grasped before high-level ones, despite the "logic" of such a statement.)

    Role Location: I E2d

20. Based on the results of the empirical data coordinated with the logical analysis, the subject will "logically" sequence instruction on unknown concepts by preceeding from familiar to unfamiliar. Some of these sequences will involve training of discriminating subclasses within a known larger class (e.g., classes of "insects"). Other sequences will involve training in generalizing to a broader class made up of known subclasses (e.g., "herbivores" from cows and deer).

    Role Location: I E2e

21. Based on the logical analysis, the subject will construct an appropriate conceptual learning criterion consisting of new examples and non-examples of each class. (Actually, they are "new" in the sense that they will now be kept out of the instructional sequence and reserved for the testing sequence.)

    Role Location: I E2f

22. In designing content for teaching frames, the subject will select widely differing examples and juxtapose these to create maximum generalization.

    Role Location: I E2g

23. In designing content for teaching frames, the subject will select the closest non-example possible for each critical attribute of the class, and juxtapose these to create fine discriminations.

    Role Location: I E2h

24. Depending on his knowledge of the verbal sophistication of the target students, the subject will adjust the verbalization of the differences between examples and non-examples and the similarity between examples and further examples of as simple a level as required.

    Role Location: I E2i

25. The subject will reserve further examples and non-examples for use in an expanded program should developmental testing demonstrate that the teaching is insufficient.

    Role Location: I E2j

26. In developmental testing, the subject will probe for student difficulties with irrelevant attributes not located by subject matter experts but proving distracting to students and with distinctions which prove too difficult.

    Role Location: I E2k

27. The subject will check each expansion of the program against the items reserved for the test to determine that he is still holding to the basic conceptual objective and is not "teaching the test."

    Role Location: I E21

28. Given that a learner population should "understand" a stated concept, the subject will perform a concept analysis and specify terminal behavioral objectives of a training program, the achievement of which will satisfy subject matter experts.

    Role Location: I E3e
    I F1i

29. Given an instructional module, the subject will state a terminal objective intended by the module.

    Role Location: I F1a

30. Given an instructional module, the subject will specify the minimum pre-learning behaviors required of learners.

    Role Location: I F1b

31. Given instructional objectives and sample measurement procedures, the subject will discriminate between appropriate and inappropriate measurement procedures.

    Role Location: I F1c
    I F2g

32. Given a specific behavioral objective, the subject will produce exemplary criterion measures for each of the categories in a given classification of criterion measures.

    Role Location: I F1j
    I F2f

33. Given specific behavioral objectives, the subject will develop appropriate criterion measures.

    Role Location: I F1k
    I F2r

34. Given a classification scheme for criterion measures and sample measures, the subject will correctly classify the sample measures.

    Role Location: I F1l
    I F2e

35. Given a specific behavioral objective and content, along with sample prerequisite skills, the subject will discriminate between essential and non-essential prerequisite skills.

    Role Location: I F1m

36. Given a specific behavioral objective and content, the subject will specify in terms of behavior and content all prerequisite skills required of learner.

    Role Location: I F1n

37. Given a discrimination task, the subject will determine whether it is one-directional or two.

    Role Location: I F1o
    I F2i

38. Given a two-way discrimination (association) task, the subject will construct appropriate criterion measures for each.

Role Location: I F1p
I F2j

39. Given a discrimination learning problem, the subject will construct criterion measures at different levels of difficulty.

Role Location: I F1q
I F2k

40. Given a criterion measure, the subject will analyze it into component behaviors.

Role Location: I F1r
I F21

41. Given a statement of prerequisites, the subject will construct a test of these, validate that test, and then determine whether such prerequisites are rational for the general population to be served. (General objectives: since not all the answers are in on instructional strategies and even when they are, they could be replaced by more elegant solutions.) All subjects should exhibit a strong tendency to keep in touch with new developments (not to be confused with research literature).

Role Location: I F1s

42. Given a set of rules or principles to be mastered, the subject will determine the kinds of components involved in the rule and to devise tests for student mastery of the components.

Role Location: I F1t

43. Given an instructional module and criterion measures, the subject will construct a decision system to monitor the progress of the learner through the module.

Role Location: II C1a

44. Given a set of instructional modules described for a population of learners, the subject will construct a decision system which identifies those modules individual learners require and may take (based on diagnostic testing).

Role Location: II C1b

45. Given the elements of a decision system with all dependencies specified, the subject will prepare PERT type analyses.

Role Location: II C1c

46. Given the specifications for an instructional module, the subject can estimate the costs of producing the module in different formats.

Role Location: II C2a

47. Given that instructional outcomes have been specified for learner populations, a subject can estimate the cost (individual, resources, organizational, developmental, etc.) of applying alternative instructional modules.

Role Location: II C2b

48. The subject will distinguish between the cost factors attributable to the functions of production in or distribution by any particular medium.

Role Location: II C2c

49. The subject will distinguish between fixed and variable media costs with respect to either the production or the distribution function.

Role Location: II C2d

50. Given educational prescriptions for a learner, the subject can select from within a system the correct diagnostic and achievement tests to be administered.

Role Location: II D1a

51. Given the results of achievement and diagnostic testing for a learner, the subject can specify the correct instructional modules and starting points for an individual learner.

    Role Location: II D1b

52. Given that a learner is progressing through a set of instructional modules, the subject can identify those points when criterion measures must be administered and select the correct criterion measures.

    Role Location: II D2a

53. Given that a selected criterion measure must be administered, the subject can administer the test and identify the correct decisions for following instruction, based on the system.

    Role Location: II D2b

54. Given a particular test, the subject can compute needed statistics (reliability, validity, error rate, required probabilities, etc.).

    Role Location: II D3a

55. Given that tests for decision purposes have been administered to a learner progressing through an instructional system, the subject can correctly compute necessary statistics (Bayesian probabilities of mastery, probability of successful completion, etc.).

    Role Location: II D3b

56. Given a set of terminal objectives, the subject can select the most effective learning strategy (chaining, etc.).

    Role Location: III C1a

57. Given a set of instructional objectives, the subject will sequence these in the most effective and/or efficient way.

    Role Location: III C1b

58. Given specific behavioral objectives, the subject will generate rules for deciding upon appropriate and inappropriate content for the instructional module.

Role Location: III C2a

59. Given specific behavioral objectives, the subject will discriminate between appropriate and inappropriate content for the objectives.

Role Location: III C2b

60. Given a specific behavioral objective, the subject will describe examples of appropriate content.

Role Location: III C2c

61. Given a set of criterion measures, the subject will try these out on appropriate students to be sure that they can be achieved by those who "know" and can't by those who don't.

Role Location: III D1a

62. Given a set of objectives and enough time to do so, the subject will design the leanest possible draft of materials to achieve the objectives and let student difficulties and queries determine the expansion at appropriate points.

Role Location: III D1b

63. Given data on differential student achievement of prerequisites, the subject will construct an individualized track leading up to the main track.

Role Location: III D1c

64. Given a set of items to be memorized by students, the subject will design an instructional system which will permit individualization of practice according to student need.

Role Location: III D1d

65. Given a set of objectives and an intended population of learners, the subject will determine an appropriate level of student achievement beyond which he will cease developmental testing. (Being reasonable, he may raise or lower this level based on practical experience and cost requirements during development.)

Role Location: III D1e

66. Given a problem-solving process to be mastered, the subject will construct exercises which require the process to be applied to already-known phenomena.

Role Location: III D1f

67. Given a terminal objective and intermediate objectives, the subject will sequence the intermediate objectives in the most effective manner.

Role Location: III D2a

68. Given a terminal objective and all necessary intermediate objectives, the subject can specify all of the dependencies among the intermediate objectives and indicate necessary branching.

Role Location: III D2b

69. Given a first draft of instructional materials, the subject will present this to an appropriate student in such a way as to maximize feedback about the materials from the student.

Role Location: III D2c
III D5e

70. The subject will handle student errors and student criticism in a completely supportive manner, encouraging rather than discouraging student feedback.

Role Location: III D2d
III D5f

71. The subject will probe student difficulties in such a way as to determine the source of the problem in the materials without giving away the appropriate response.

Role Location: III D2e

72. Given a tape or transcript of student feedback, the subject will make changes in the instructional materials which retain the original objectives but reduce the likelihood of error in in the next student.

Role Location: III D2f

73. Given a set of discriminations to be learned, the subject will arrange them according to least potential for interference, if the learning ability of the students makes this possible.

Role Location: III D2g

74. Given a set of component behaviors, the subject will determine whether a logical hierarchy is implied in sequencing these for instruction.

Role Location: III D2h

75. Given a logical hierarchy, the subject will validate its existence by appropriate tryouts with students.

Role Location: III D2i

76. Given a set of objectives with no predetermined sequence, the subject will select an appropriate sequence if time does not permit further exploration with students.

Role Location: III D2j

77. If time permits further exploration, the subject will determine which sequencing appeals most to students by observing their inquiry behaviors.

Role Location: III D2k

78. Given a principle to be mastered, the subject will sequence it in the instruction in such a way that almost all elements in it have been mastered. (No data exist on how many elements might be mastered concurrently in complex principles in physics or grammar -- but we would say not many!)

Role Location: III D2l

79. Given an instructional objective and a set of reinforcers, the subject will discriminate between the appropriate and inappropriate reinforcers, based on the characteristics of a population of learners.

Role Location: III D3a

80. Given the results of the administration of instructional modules to a group of learners, the subject will develop more effective reinforcement schedules.

Role Location: III D3b

81. Given an instructional objective, the subject will specify effective and/or efficient reinforcers, based on the relevant characteristics of a population of learners.

Role Location: III D3c

82. Given a student or group of students going through a draft of instructional materials, the subject will observe all signs of boredom and disinterest and will probe for student attitude in post-interviews.

Role Location: III D3d
III D5g

83. Given an instructional objective, and a set of instructional procedures using different media, the subject will discriminate between more appropriate and less appropriate uses of media.

Role Location: III D4a

84. The subject will select an appropriate medium, given the demands imposed by the task analysis in interaction with the available funds, the on-going system, and the motivational requirements.

Role Location: III D4b

85. Given the cheapest medium within which the instruction might operate (e.g., straight print, for instance), the subject will prepare alternate suggestions for increasing the motivational impact.

Role Location: III D4c

86. The subject will make first-draft adaptations of timing, vocabulary and sentence length, and question difficulty appropriate to information presentation in an oral medium.

Role Location: III D4d

87. The subject will select instructional media which optimize the benefit of all resources, i.e., cost of media, cost of instructional staff, and cost of student time committed to the instructional effort.

Role Location: III D4e

88. The subject will identify media by type (and combination of types) of presentation capabilities (stimulus control), i.e., discriminations or generalizations on the basis of three-dimensional properties, sound, motion, tactile, color, taste, or verbal description. As used here, 3-D refers to all aspects of spatial and relative location. Tactile includes texture, weight or mass, and psychomotor stimuli.

Role Location: III D4f

89. Given an operationally specific objective, the subject list the available instructional media capable of presenting the type of stimulus (or combination of stimuli) to which the student must attend, as specified by the indicator behavior and conditions of the specific objective.

Role Location: III D4g
III D6a

90. Having listed available media for a given specific objective, the subject will consider all factors and select for the first draft trial the medium (or combination of media) which, on a best judgement basis, has a reasonable probability of conveying the instructional intent.

Role Location: III D4h

91. During developmental testing of instruction, the subject will employ a "lean programming" rationale with respect to mediated instruction, i.e., utilize prompting and other techniques appropriate to the media selected in order to direct attending-to behaviors to the salient aspect according to instructional intent.

Role Location: III D4i

92. The subject will recognize the point at which further efforts to "prompt" the initial selection of media are uneconomical during developmental testing and, at that point, shift to the "mere representative" listed media along the cost continuum.

Role Location: III D4j

93. The subject will evaluate the level of "media sophistication" of the SME, basing initial media selection on this level -- as one factor of cost -- in view of the time demand required as the SME interacts with media production personnel.

Role Location: III D4k

94. The subject will control the understandable tendencies of graphics personnel to create "mini-works-of-art" when the limits of the instructional situation, i.e., first draft trial or small number of students, do not justify an unreasonable expenditure of talent and production resources.

Role Location: III D4l

95. Given a specific behavioral objective and content, the subject will construct elicitors, cues, problems, etc.

Role Location: III D5a

96. Given a specific behavioral objective and content, the subject will describe appropriate responses.

Role Location: III D5b

97. Given the essential specifications for instructional modules (objectives, content, criterion measures, elicitors, cues, problems, etc.), a subject will write effective instructional frames.

Role Location: III D5c

98. Given a summary data from the administration of an instructional module, the subject will identify ineffective frames, sequencing, stimuli, etc.

Role Location: III D5d

## PART 3 - OBJECTIVES FOR EDUCATIONAL EVALUATION

### A. Rationale

As mentioned earlier, the guiding principle of the proposed consortium is that training will be provided in terms of desired terminal behaviors rather than in terms of courses taken or time spent. Prerequisite to the training of educational evaluators is the need to define precisely the desired instructional objectives for evaluation. The next step in the training process is to develop appropriate training packages to achieve these objectives. Implementation of the training packages on an individual basis should result in the desired terminal behaviors. In this conceptual paper we apply this sequence to training educational evaluators. Again, the steps in our proposal for training evaluators are as follows:

1. Determine the instructional objectives for training educational evaluators.
2. Translate these into a set of terminal behaviors.
3. Develop or select appropriate training modules to accomplish these terminal behaviors.
4. Conduct diagnosis of current competencies and implement appropriate training packages on an individual basis.

As part of the design proposal, we present in this paper hierarchies of evaluation behaviors together with a rationale for the creation of this set of behaviors. These hierarchies include both program evaluation and student evaluation. Examples of possible training modules are also presented although the final definition of competencies and training will be created during the implementation stage.

The identification of evaluation behaviors can be accomplished in several ways. To identify appropriate instructional objectives for training persons skilled in program and student evaluation we have conducted self-analyses among staff members, and selected sets of objectives from other published sources. It appears that any of these sources, or more likely, some combination thereof, might serve as the basis of the evaluation training implementation stage. Certainly further information from the Schalock study will help to delineate the tasks and functions of evaluators more clearly. For the moment, it seems best to present typical objectives for evaluators leaving the specifics to the implementation stage.

The first part of this chapter is concerned with program evaluation and draws upon instructional objectives as derived from Welch (1970), and Coller (1969). The section on student and instructional evaluation is largely from Terwilliger (1970).

## B. A Hierarchy of Program Evaluation Objectives

Evaluation is the gathering of information for the purpose of making effective decisions. In education, these decisions are usually concerned with programs, students, and instruction. Traditionally, educational evaluation has focused on students and instruction, but in recent years a growing need for program evaluation has developed. Essential to successful evaluation in all three areas is the need to determine the kinds of decisions that have to be made.

Decisions regarding programs usually are made in response to one of three questions. How can we improve the program (formative)?

Should we adopt a program (summative)? Have we invested our resources wisely (administrative)? The influx of federal monies into local education has stimulated considerable interest in program evaluation. New curriculum projects, the various title projects of the USOE, and the increased concern of local school districts have created a need for new strategies of program evaluation. Qualified personnel in this area is one of the primary concerns of the proposed Education Training Center.

An analysis of the skills possessed by several practicing evaluators was conducted by the author. These skills appeared to arrange themselves along a hierarchy of complexity and are presented in this fashion. This is not to suggest that there is a hard and fast set of prerequisite skills for each level, but that experience and training are more in demand as one moves from bottom to top. Within each level of educational objectives is presented a sample of instructional objectives. A format and several examples of objectives translated into terminal behaviors are found in Section E.

For each of the selected instructional objectives a training module will be developed. The mechanism by which these objectives are brought together will be derived from either a role definition or a scientifically identifiable training program.

TABLE 1

PROGRAM EVALUATION

HIERARCHY OF EDUCATIONAL OBJECTIVES (Welch, 1970)

| Level | |
|---|---|
| I. | Develops theory and conducts research on the evaluation process. |
| II. | Designs specific evaluation studies. |
| III. | Conducts audits of planned and ongoing evaluations. |
| IV. | Interprets evaluation results. |
| V. | Manages or administers an evaluation study. |
| VI. | Communicates to decision makers appropriate information |
| VII. | Analyzes and summarizes the meaning of data. |
| VIII. | Develops data-gathering techniques. |
| IX. | Selects appropriate data-gathering techniques for specific evaluation problems. |
| X. | Writes and utilizes behavioral objectives. |
| XI. | Administers data-gathering process. |
| XII. | Processes data including key punching, test scoring, etc. |

TABLE 2

PROGRAM EVALUATION

SAMPLE INSTRUCTIONAL BEHAVIORS FOR EACH LEVEL

Level

I. THEORY AND RESEARCH
- A. Is able to create a new evaluation technique.
- B. Constructs a theory of curriculum evaluation.
- C. Quantifies the decision-making process.

II. DESIGN
- A. Develops strategy to evaluate a specific course.
- B. Writes an acceptable proposal in response to an evaluation problem.

III. AUDIT
- A. Is able to identify biases in a proposed evaluation design.
- B. Can determine if a chosen strategy is appropriate to its task.

IV. INTERPRETATION
- A. Is able to select data from a study pertinent to a given decision.
- B. Can differentiate between descriptive and judgemental data.

V. MANAGEMENT
- A. Operates a study within budget levels.
- B. Can select personnel able to perform the necessary evaluation tasks.

VI. COMMUNICATION
- A. Writes understandable reports.
- B. Presents data in a concise and clear manner.
- C. Establishes liason between program developer and program evaluator.

VII. ANALYSIS
- A. Selects appropriate statistical treatment.
- B. Can use computers and/or calculators.
- C. Computes item analyses.
- D. Can write computer programs.

VIII. DEVELOPMENT
- A. Can write valid, reliable, and useable attitude measure.
- B. Uses item analysis results to revise an achievement test.
- C. Selects or creates a non-contaminative evaluation technique, i.e., unobtrusive measure.

IX. SELECTION
- A. Given a set of achievement tests, can select the one most suited to the specific evaluation problem.
- B. Knows the existence of and can use the various test review services, e.g., NCME reviews.

X. BEHAVIORAL OBJECTIVES
- A. Given a set of objectives, is able to identify which are written in behavioral terms.
- B. Can write or identify the behavioral objectives for the program being evaluated.

XI. IMPLEMENTATION
- A. Can conduct interviews for evaluation purposes without contaminating the results.
- B. Administers group attitude questionnaires.

XII. PROCESS
- A. Uses scoring keys to correct achievement tests.
- B. Can key punch computer cards.
- C. Submits to and uses canned computer programs.

A plan for translating the instructional objectives into desired terminal behaviors has been agreed upon by the UMC consortium members. Several examples of this plan are listed for training evaluators in Section E. However, because final agreement on the specific objectives will not be made until the implementation phase, only a few examples are presented.

C. A Taxonomy of Tasks in an Educational Evaluation Facilitation and Coordination System (from Coller, 1970)

In the pages that follow is listed another scheme of evaluation objectives. The format of the set of tasks derived by Coller (1970) is in slightly different format from that proposed by Welch yet the basic purpose is the same--to present a set of objectives considered important as training goals for evaluators.

The members of the consortium that will be developing modules and providing training will draw upon this objective pool for their content. For the present no attempt has been made to equate these objectives to the various levels of the Welch hierarchy. However, such a selection will be made prior to the development of the training modules.

1. Developing Supportive Climates for Evaluation
   - 1.1 Developing alternative strategies based upon particular characterisitics of select members or segments of the educational milieu in order to modify reactions toward evaluation
   - 1.2 Assessing the existing attitudes of selected members of the educational milieu towards evaluation
   - 1.3 Becoming thoroughly aware of the individual's or group's unique characteristics
   - 1.4 Establishing trust and rapport with all members of the educational milieu, but particularly with participants of an evaluation effort
   - 1.5 Identifying available neoteric evaluation strategies and techniques
   - 1.6 Making select members of the educational milieu aware of the discrepancies between the observed and the intended
   - 1.7 Demonstrating, in other ways, the efficacy of evaluation
   - 1.8 Identifying and reducing any inhibitions toward evaluation
   - 1.9 Instilling within the educational community an awareness of the need for evaluation
   - 1.10 Instilling within the educational community a demand for evaluation
   - 1.11 Reinforcing positive attitudes toward evaluation
   - 1.12 Stating clearly the purposes of an evaluation and the role of the evaluator
   - 1.13 Indicating clearly, when appropriate, the types of information that will result from an evaluation and to whom it will be made available
   - 1.14 Establishing clarly operational procedures and mutual rules of "etiquette" with participants of an evaluation effort

1.15 Maintaining open communication channels among the participants of an evaluation effort
1.16 Thanking the participants of an evaluation effort for their cooperation
1.17 Crediting, when appropriate, selected participants for their cooperation and aid in bringing an evaluation effort to a conclusion
1.18 Providing, when appropriate, feedback related to the outcomes of an evaluation effort to participants of the evaluation
1.19 Inviting and encouraging select members of the educational milieu to participate actively in evaluation efforts
1.20 Supporting the efforts of persons initiating and/or engaging in evaluation efforts
1.21 Consulting with and supporting the efforts of professionals and para-professionals who are attempting to develop a supportive climate for evaluation

2. Focusing an Evaluation: I. Selecting Decisions Situations
2.1 Training select members of the educational milieu to become sensitive to decision stimuli (including "institutional presses")
2.2 Responding to decision stimuli (including "institutional presses")
2.3 Reinforcing those individuals sensitive to decision st stimuli
2.4 Identifying decision situations of interest and importance to select members of the educational milieu
2.5 Determining the desirability of collecting evaluative information for each of the identified decision situations
2.6 Identifying the major level(s) of decision-making and of decision-makers to be served
2.7 Identifying the decision-making process as it operates in a given setting
2.8 Determining meaningfulness and feasibility of collecting evaluative information for each of the identified decision situations
2.9 Establishing priorities for the collection of evaluative information
2.10 Selecting tentative, desirable, meaningful, and feasible decision situations for evaluation

3. Focusing an Evaluation: II. Selecting Evaluation Strategies and Developing Evaluation Plans
3.1 Establishing criteria for decision-making regarding expected outcomes
3.2 Projecting the decision situations to be served in terms of their locus, focus, criticality, timing, and composition of alternatives
3.3 Becoming knowledgeable about relevant aspects of settings, conditions, and/or contexts within which the evaluation will occur
3.4 Projecting the decision situations to be served in terms of political, social, institutional, and situational constraints

3.5 Defining policies and limits within which the evaluation must operate
3.6 Making explicit and clarifying project assumptions
3.7 Requesting pertinent information from surveillance specialists and retrieval centers
3.8 Reviewing research literature concerning similar projects in order to: (1) verify assumptions, (2) uncover sources of possible incidental gains or unwanted side effects, and (3) make as uniform as possible the use of tests and testing procedures
3.9 Visiting pertinent persons and places to acquire updated information
3.10 Developing the rationale and objectives for the decisions situation
3.11 Identifying or formulating the basic question and/or hypothesis of the evaluation
3.12 Establishing premises which will guide the evaluation
3.13 Determing if the evaluation goals are formative, summative, or both
3.14 Determining the level of generalization for the evaluation
3.15 Identifying, when appropriate, the pre-experimental, experimental, or quasi-experimental research designs to be used in the evaluation
3.16 Identifying available evaluation strategies
3.17 Selecting an appropriate evaluation strategy
3.18 Utilizing, when necessary, the services of evaluation specialists to select evaluation strategies and develop evaluation plans
3.19 Developing, if necessary, evaluation strategies
3.20 Identifying segments of the educational community to be affected by the intended transactions and outcomes
3.21 Detailing, when appropriate, intendent antecedents, transactions, and outcomes
3.22 Re-structuring, when necessary, intentions (objectives) into observable (measurable or describable) outcomes
3.23 Listing, if appropriate, the potential existence of contingencies between antecedents, transactions, and outcomes
3.24 Listing, if appropriate, the potential congruence between intents and observations
3.25 Making explicit and clarifying standards for use in the judgment of alternatives
3.26 Identifying and clarifying judgments required of the evaluator
3.27 Identifying segments of the educational milieu from which judgments will be collected
3.28 Estimating, when appropriate, which data-gathering techniques will be utilized in the collection of observables
3.29 Identifying the sample
3.30 Relating the project or program to other efforts or experiences of others who have coped with similar or related areas, and showing how the project utilizes, builds upon, extends, revises, or adapts to existing knowledge
3.31 Coordinating efforts to develop evaluation plans
3.32 Smoothing the efforts of others to develop evaluation plans

4. Organizing the Evaluation and Developing Managerial Plans
   - 4.1 Reviewing and becoming familiar with the objectives of the evaluation effort
   - 4.2 Identifying those program and/or project components which must be completed in order to achieve the evaluation objectives
   - 4.3 Reviewing and becoming familiar with issues of criticality and timing
   - 4.4 Identifying "milestones" and developing a planning structure for the program or project
   - 4.5 Developing planning structures for individual projects or components
   - 4.6 Developing plans for selecting or developing data-gathering techniques
   - 4.7 Developing sampling plans
   - 4.8 Developing plans for coding data
   - 4.9 Developing plans for collecting data
   - 4.10 Developing plans for preparing raw data
   - 4.11 Developing plans for treating data
   - 4.12 Developing plans for interpreting and judging outcomes
   - 4.13 Developing plans for reporting data
   - 4.14 Developing plans for storing and disseminating information
   - 4.15 Developing plans for information utilization
   - 4.16 Determining the dependency relationship existing among the identified project components
   - 4.17 Arranging the components of the evaluation in a network according to a plan
   - 4.18 Determining how long it will take to complete the program plan
   - 4.19 Determining the amount and type of resources required to complete the plan
   - 4.20 Modifying the plans according to known restraints
   - 4.21 Determining local resources available for the evaluation
   - 4.22 Defining the manpower and material needs for the evaluation
   - 4.23 Constructing an evaluation budget for the project
   - 4.24 Coordinating efforts to develop managerial plans
   - 4.25 Smoothing efforts to develop managerial plans

5. Assessing, Modifying, and Selecting Evaluation and Managerial Plans
   - 5.1 Reviewing evaluation and managerial plans for communication value, i.e., are the plans clear and seductive
   - 5.2 Determining if the scope of the evaluation has been stated explicitly
   - 5.3 Determining, when necessary, if the format of the proposed plans is appropriate for the receiving agency
   - 5.4 Learning the project rationale, objectives, and operational procedures
   - 5.5 Determining the audiences, the decision-makers to be served, and the nature of the implementing agency

5.6 Using effectively subject area or technical specialists whenever necessary to review the evaluation and managerial plans
5.7 Determining the relevance of the proposed evaluation plan to the identified decision situations
5.8 Determining the relevance of the proposed managerial plan to the proposed evaluation plan
5.9 Determining the legal status of the proposed evaluation and managerial plans relative to the context within which they are to be implemented
5.10 Determining the congruence of the evaluation and managerial plans with the value systems of the context within which they are to be implemented
5.11 Determining if the evaluation plan is within the purview of the agency charged with the implementation
5.12 Determining the compatibility of the evaluation and managerial plans with the value system(s), i.e., purposes and goals of the implementing agency
5.13 Determining the impact of the evaluation and managerial plans on other components (sub-systems) of the system and on the weights and interrelationships of these system elements
5.14 Determining the practicality of the evaluation and managerial plans in terms of achieving its stated purposes (end-products)
5.15 Determining the relative desirability of the evaluation and managerial plans (in comparison with other plans) in terms of the ratio of necessary inputs (costs) to expected outputs (effectiveness)
5.16 Consulting with clients in order to review and/or modify evaluation and managerial plans
5.17 Modifying the evaluation and managerial plans in terms of the outcomes of the assessment
5.18 Coordinating assessment efforts
5.19 Smoothing assessment efforts
5.20 Submitting the proposed plans for approval and funding

6. Selecting, Modifying, or Developing Data-Gathering Techniques
6.1 Reviewing and becoming familiar with the intents of an evaluation study
6.2 Stating the purposes for which data-gathering techniques are to be used
6.3 Stating explicitly the objectives for which data-gathering techniques are to be used
6.4 Utilizing, when necessary, the services of test and measurement specialists
6.5 Reviewing resource materials related to similar projects to uncover suitable data-gathering techniques
6.6 Securing and maintaining copies of data-gathering techniques

6.7 Utilizing, when necessary, the services of surveillance specialists
6.8 Developing criteria for selecing the most suitable available data-gathering techniques
6.9 Identifying pertinent techniques for which evidence indicates effectiveness
6.10 Selecting suitable data-gathering techniques
6.11 Modifying, field-testing, and revising identified data-gathering techniques
6.12 Developing specification for constructing data-gathering techniques if no existing data-gathering techniques are appropriate
6.13 Developing, pilot testing, and revising new data-gathering techniques
6.14 Collecting, when appropriate, reliability and validity information
6.15 Collecting, when necessary, normative information for use as standards
6.16 Preparing administrative and scoring manuals for the newly modified or developed data-gathering techniques
6.17 Training personnel in the administration of various data-gathering techniques
6.18 Consulting with clients regarding available data-gathering techniques
6.19 Advising clients regarding the development, validation, and norming of various data-gathering techniques
6.20 Aiding select members of the educational milieu in the application of sound tests and measurement policies, programs, and practices
6.21 Persuading others to cooperate in the development, validation, and norming of specific data-gathering techniques
6.22 Obtaining resources and resource personnel to facilitate the development, validation, and norming of specific data-gathering techniques
6.23 Coordinating efforts to select, develop, modify, validate, and/or norm data-gathering techniques
6.24 Smoothing the efforts of others to select, modify, develop, validate, or norm data-gathering techniques
6.25 Researching new data-gathering techniques

7. Collecting Data
7.1 Specifying information needs clearly and concisely
7.2 Identifying information sources (populations and individuals) for the collection of data
7.3 Identifying information environments for collecting data
7.4 Specifying methods to be used in collecting data
7.5 Specifying sampling procedures
7.6 Specifying the schedule for data collection
7.7 Reviewing the sampling plan and schedule with relevant others for appropriateness and congruence with other on-going programs

7.8 Utilizing, when necessary, the services of data collection specilaists
7.9 Preparing the sample population and relevant others for data collection
7.10 Field-testing data collection methods
7.11 Training personnel to collect and record data
7.12 Informing personnel of the rules of "etiquette" for collecting data
7.13 Administering evaluative data-gathering techniques and recording the data
7.14 Reviewing and modifying, when necessary, data collecting plans and methods
7.15 Obtaining resources and resource personnel for data collection efforts
7.16 Persuading others to cooperate in the collections and recording of data
7.17 Coordinating efforts to collect data
7.18 Smoothing the efforts of others to collect and record data
7.19 Researching data collection techniques

8. Data Processing: I. Preparing "Raw" Data
8.1 Providing specifications for the scoring and/or classification of data
8.2 Becoming familiar with the data, the intended analytic process, and available computer programs
8.3 Providing formats for coding data which are compatible with available computer programs or analytic procedures
8.4 Training personnel in response interpretation
8.5 Training personnel to operate mechanical scoring units
8.6 Scoring and/or classifying responses obtained from the administration of data-gathering techniques
8.7 Utilizing, when necessary, the services of professional scoring services
8.8 Transcribing memorial or taped data when necessary
8.9 Training personnel to operate machines related to the preparation of computer data cards
8.10 Transferring raw data onto computer cards, tabulation sheets, or other data storage systems
8.11 Utilizing, when necessary, the services of data processing assistants
8.12 Providing for data storage, management and retrieval
8.13 Utilizing, when necessary, the services fo data processing specialists, and test measurement specialists
8.14 Informing relevant members of the educational milieu regarding the types of raw data available for heuristic purposes
8.15 Coordinating data processing activities
8.16 Disseminating raw data upon the request of appropriate officials
8.17 Consulting with clients regarding the preparation of raw data

8.18 Obtaining resources and resource personnel for preparing raw data
8.19 Smoothing the efforts of others to prepare raw data

9. Data Processing: II. Treating Data (Analysis)
9.1 Reviewing the objectives of the evaluation
9.2 Determining the level of sophistication required by decision-makers
9.3 Reviewing the evaluation design actually employed
9.4 Reviewing the sampling procedures actually employed
9.5 Determining the nature of the data collected
9.6 Determining the desired levels of statistical precision
9.7 Reviewing the research literature for new statistical procedures
9.8 Utilizing, when necessary, the services of statistical specialists
9.9 Becoming familiar with available computer programs
9.10 Selecting the analytical procedures
9.11 Using existing computer programs
9.12 Writing new computer programs when necessary
9.13 Utilizing, when necessary, the services or program writers
9.14 Designating a means for performing the analysis
9.15 Developing a library of cookbook-like statistical forms and procedures
9.16 Performing the statistical computations
9.17 Treating data to test hypothesis determining relationships and/or to answer basic questions
9.18 Obtaining resources and resource personnel to facilitate the analytical process
9.19 Utilizing, when necessary, the services of computer and computer-related equipment operators
9.20 Coordinating efforts to treat the data
9.21 Smoothing the clients of others to treat data
9.22 Consulting with clients regarding statistical procedures
9.23 Researching new analytical procedures
9.24 Producing computational documentation when appropriate
9.25 Advising clients regarding available computer programs
9.26 Consulting with clients regarding computer programming and operations
9.27 Maintaining a library of computer programs and their descriptions
9.28 Determining the types of computer programs which will be demanded in the future
9.29 Disseminating information pertaining to the acquisition and/or modification of computer programs

10. Interpreting and Judging Outcomes
10.1 Reviewing and objectives of the evaluation
10.2 Becoming thoroughly familiar with the transactions of the evaluation; the basic questions asked; the criteria, model, procedure, and techniques employed; and the data collected

10.3 Verifying the statistical analysis and the procedures used to collect and process the data
10.4 Comparing observed antecedents, transactions and outcomes with intended antecedents, transactions, and outcomes
10.5 Describing points of congruence and incongruence
10.6 Making inferences about contingencies among transactions and outcomes
10.7 Interpreting the results of the evaluation program in terms of given criteria
10.8 Establishing relationships and distinguishing between those outcomes that result from treatment application of those contingent upon antecedent conditions
10.9 Utilizing, when necessary, the services of evaluation, statistical, and subject area specialists in order to interpret the data
10.10 Determining if any unintended outcomes occurred
10.11 Identifying the ways in which segments of the educational community, i.e., the institution, its staff, students, and community, etc., were affected by the observed transactions and outcomes
10.12 Comparing generalizations from the literature and theory with those drawn from the evaluation outcomes
10.13 Reviewing the judgments required of the evaluator
10.14 Obtaining judgments of concerning the outcomes of the evaluation from relevant members of the educational milieu
10.15 Judging the "worth" of the outcomes of the evaluation
10.16 Rendering judgments regarding the worth of alternative strategies as employed in the evaluation
10.17 Rendering judgments as to the significance of the observed transactions for various segments of the educational milieu
10.18 Judging if the unintentional outcomes are unwanted side effects or incidental gains
10.19 Rendering judgments regarding the worth and relevance of data-gathering techniques as used in the evaluation
10.20 Rendering judgments as to the overall quality of the evaluation effort
10.21 Recommending, if appropriate, future modifications for the data-gathering techniques
10.22 Identifying and recommending alternative strategies
10.23 Identifying and recommending procedures to control or reduce unwanted side effects
10.24 Identifying and recommending procedures to control or enhance incidental gains
10.25 Providing counsel to relevant members of the educational milieu regarding the interpretations and implications of the judgments rendered
10.26 Consulting with clients regarding interpretation and judging outcomes
10.27 Obtaining resources and resource personnel to facilitate the interpreting and judging of the outcomes

10.28 Coordinating efforts to interpret and judge outcomes
10.29 Smoothing efforts to interpret and judge outcomes

11. Reporting Outcomes
- 11.1 Becoming familiar with available means for reporting outcomes
- 11.2 Researching new methods for reporting outcomes
- 11.3 Specifying the format for evaluation reports
- 11.4 Becoming aware of the professional understandings of those who are to receive the evaluation results
- 11.5 Scheduling the reporting of outcomes
- 11.6 Specifying means for reporting the outcomes of the evaluation to relevant audiences
- 11.7 Coordinating efforts to prepare a report
- 11.8 Preparing a report of the evaluation that will be understandable to the public. It will serve and include meaningful terms, tables, charts, graphs, illustrations, and answers to the basic questions of the evaluation
- 11.9 Describing the intended antecedents, transactions, and outcomes
- 11.10 Describing processes and procedures by which evaluative data were gathered and judgments rendered
- 11.11 Describing the observed antecedents, transactions, and outcomes
- 11.12 Describing criteria
- 11.13 Describing explicitly unintentional outcomes
- 11.14 Detailing rendered judgments
- 11.15 Reporting the limitations of the evaluation
- 11.16 Utilizing when necessary, the services of reporting specialists in order to report the outcomes
- 11.17 Obtaining resources and resource personnel to facilitate the reporting of outcomes
- 11.18 Coordinating efforts to produce a report
- 11.19 Preparing findings and recommendations to the decision-makers in an understandable manner
- 11.20 Obtaining the decision-makers' reactions to the report
- 11.21 Providing evaluation abstracts or summaries for presentation to specific groups
- 11.22 Packaging the outcomes of the evaluation for purposes of presentation to appropriate public
- 11.23 Obtaining reactions to the presentations
- 11.24 Consulting with clients regarding the reporting of outcomes
- 11.25 Smoothing the efforts of others to prepare and produce reports
- 11.26 Reporting subsequent modifications in transactions and their observed outcomes

12. Information Processing and Dissemination
- 12.1 Providing specifications for the coding and cataloging of evaluative information and processed data
- 12.2 Training personnel in information processing and dissemination

12.3 Detailing the scope of information processing and dissemination activities performed by the given agency
12.4 Identifying the potential sources of evaluative information and processed data
12.5 Requesting evaluation information and processed data from likely sources
12.6 Collecting evaluative information and processed data
12.7 Visiting pertinent persons and places to acquire updated evaluation information
12.8 Utilizing, if necessary, the services of surveillance and subject area specialists
12.9 Conducting conferences regarding specific evaluation-related issues of current importance
12.10 Requesting select members of the EM to prepare reviews of specific evaluation areas
12.11 Coding and cataloging evaluation information and processed data
12.12 Utilizing the services of information processing specialists
12.13 Transferring evaluation information into information storage systems
12.14 Providing for information storage, management, and retrieval
12.15 Implementing and maintaining a data bank
12.16 Becoming familiar with available information media methods for reporting information
12.17 Developing, if necessary, new reporting styles
12.18 Specifying the format for reporting evaluation information
12.19 Scheduling the reporting of evaluation information
12.20 Reporting evaluation information in a highly communicable fashion
12.21 Providing evaluation abstracts or summaries
12.22 Utilizing, if necessary, the services of reporting specialists
12.23 Identifying potential recipients of evaluative information and processed data
12.24 Informing relevant members of the educational milieu regarding the types of evaluation information which may be requested
12.25 Communicating to potential users regarding the availability and use of the data bank
12.26 Packaging the information for purposes of presentation to appropriate publics
12.27 Disseminating evaluative information and processed data
12.28 Consulting with clients regarding the retrieval of evaluative information and processed data
12.29 Coordinating evaluation information processing and dissemination activities
12.30 Coordinating data bank activities
12.31 Smoothing the efforts of others to acquire specific evaluation information and processed data

12.32 Obtaining resources and resource personnel to process and disseminate evaluation information and processed data
12.33 Obtaining reactions from relevant members of the educational milieu regarding the reporting of evaluation information
12.34 Surveying select members of the educational milieu regarding their evaluation information needs
12.35 Determining future evaluation information needs

13. Decision Making: The Utilization of Information
13.1 Becoming familiar with the contents of an evaluation report
13.2 Assessing the perspicacity of the evaluation report
13.3 Using effectively select members of the educational milieu to assess the perspicacity of the evaluation report
13.4 Determining the nature of the institutions involved in the decision solution
13.5 Reviewing the major level(s) of decision-making to be served
13.6 Reviewing the decision-making process as it operates in a given setting
13.7 Developing taxonomies of educational decisions
13.8 Identifying decision-making criteria
13.9 Selecting decision-making criteria
13.10 Identifying decision-making models
13.11 Developing, if necessary, a decision-making model
13.12 Selecting a decision-making model
13.13 Identifying and proposing alternative decision solutions
13.14 Assessing the decision solution for its relevance, legality, congruence, relatedness, compatability, impact, practicality, and relative desirability
13.15 Utilizing effectively select members of the educational milieu to assess the proposed decision solution
13.16 Selecting desirable and feasible decision solutions
13.17 Reinforcing those individuals who produce "rationale" decision solutions
13.18 Reinforcing those individuals who produce "creative and feasible" decision solutions
13.19 Training select members of the educational milieu in decision-making strategies
13.20 Inviting and encouraging select members of the educational milieu to participate actively in decision-making processes
13.21 Coordinating the decision-making process
13.22 Providing directives, guidelines, and/or other needed assistance to decision-makers for purposes of improving the usefulness of transmitted data in the decision-making process
13.23 Smoothing the efforts of decision-makers to derive and select decision solutions

13.24 Obtaining reactions to the decision solutions from relevant members of the educational milieu
13.25 Assisting the decision-maker in formulating new questions for future evaluation based upon the original evaluation findings
13.26 Using evaluation findings as the basis for discussion in teacher in-service training sessions
13.27 Conducting sensitivity sessions to induce selected members of the educational milieu to make use of the evaluative information in decision solutions
13.28 Suggesting to decision-makers techniques by which they themselves can evaluate and modify behavior

14. Performing Activities and Research Related to Evaluation and to the Development, Diffusion, and Adoption of Evaluation
14.1 Defining explicitly the meanings of evaluation-based terms
14.2 Developing and testing evaluation models
14.3 Performing research related to new methodological techniques
14.4 Drafting plans for constructing idealized evaluation strategies and techniques for usage in select settings of the educational milieu
14.5 Constructing idealized evaluation strategies and techniques for usage in select settings of the educational milieu
14.6 Integrating the components of the strategies and techniques into operating systems for usage in select settings of the educational milieu
14.7 Developing procedures for creating widespread awareness of the neoteric evaluation strategies and techniques
14.8 Developing situations by which individuals can examine and assess operating qualities of the neoteric evaluation strategies and techniques
14.9 Researching procedures for the training of local personnel to manage, operate, service, and utilize neoteric evaluation strategies and techniques
14.10 Developing situations for the trial use of the evaluation strategies and techniques
14.11 Modifying the neoteric evaluation strategies and techniques to fit the particular circumstances of the adopting institution
14.12 Performing studies to assure the assimilation of the evaluation strategies and techniques by the adopting institution
14.13 Conducting longitudinal studies to determine effects of specific variables over time
Performing experimental research on some of the substantive areas being evaluated
14.15 Collecting standards of all kinds
14.16 Developing a taxonomy for standards

14.17 Performing case studies or other types of research to learn more about the nature of children involved in programs being evaluated
14.18 Conducting simulation studies and predictive studies
14.19 Conducting surveys related to educational needs, and uses and abuses of evaluation
14.20 Determining the applicability of various data-gathering techniques for special populations
14.21 Comparing alternative strategies and techniques for instilling in select members of the educational milieu an awareness of a need for evaluation
14.22 Comparing alternative strategies for instilling in select members of the educational milieu a demand for evaluation
14.23 Determining attitudes toward evaluation and readiness for change
14.24 Coordinating research and activities relating to the development, diffusion and adoption of evaluation
14.25 Coordinating evaluation based research in general
14.26 Smoothing general research activities and activities related to development, diffusion and adoption of evaluation

15. Administering and Coordinating Activities in an Evaluation Facilitation and Coordination System
15.1 Stating explicitly the broad purposes of an evaluation and/or facilitation entity
15.2 Developing specific policies and general guidelines for the operation of an evaluation and or facilitation entity
15.3 Identifying and assessing alternative objectives as possible goals for an evaluation and or facilitation entity
15.4 Defining criteria for selecting objectives for an evaluation and/or facilitation entity
15.5 Selecting and assigning priorities to objectives for an evaluation and/or facilitation entity
15.6 Stimulating and assisting in periodic evaluation, reflection, and revision of purposes and or objectives
15.7 Defining staff and resource requirements for operating an evaluation and/or facilitation entity
15.8 Developing plans to meet staff and resource requirements
15.9 Developing job descriptions
15.10 Constructing, securing, and managing budgets
15.11 Developing policies and procedures for the selection, assignment, retention, dismissal, promotion, and in-service growth of personnel
15.12 Establishing criteria for evaluating the on-the-job performance of personnel
15.13 Developing policies and techniques for evaluating on-the-job performance of personnel
15.14 Reviewing all evaluation designs, instruments, and reports before they are used or released for distribution
15.15 Identifying sources, i.e., foundations or agencies, which have indicated an interest in supporting programs or projects similar in kind to the submitting evaluation and/or facilitation entity

15.16 Making an informal contact with the agency to which the proposal will be submitted
15.17 Developing an overall managerial plan for an evaluation and/or facilitation entity
15.18 Supervising the training, research, facilitation and coordination services performed by staff members
15.19 Arranging for in-service training of the staff
15.20 Organizing the tasks within the entity in order to utilize the unique talents of each member
15.21 Maintaining conditions conducive to high morale and job efficiency
15.22 Arranging for an independent evaluation of the activities of an evaluation and/or facilitation entity
15.23 Coordinating evaluation, facilitation and/or coordinating activities within the evaluating group
15.24 Coordinating evaluation, facilitation and/or coordinating activities within the school
15.25 Coordinating evaluation, facilitation and/or coordinating activities within the district
15.26 Coordinating evaluation, facilitation and/or coordinating activities between districts
15.27 Coordinating evaluation, facilitation and/or coordinating activities within a cooperative multi-district unit
15.28 Coordinating evaluation, facilitation and/or coordinating activities between cooperative multi-district units
15.29 Coordinating evaluation, facilitation and/or coordinating activities within the state
15.30 Smoothing the administrating and coordinating efforts of others

16. Providing Facilitation and Coordination Services
16.1 Surveying the training needs of select members of the educational milieu
16.2 Developing instructional objectives, plans, aids, and materials for training
16.3 Training select members of the educational milieu in tasks associated with their roles
16.4 Training select members of the educational milieu in tasks related to evaluation facilitation and/or coordination
16.5 Obtaining from select members of the educational milieu reactions to training
16.6 Surveying the service needs of select members of the educational milieu
16.7 Facilitating, generally, the efforts of select members of the educational milieu to undertake evaluations
16.8 Assisting select members of the educational milieu to develop objectives
16.9 Reducing impediments to evaluations

16.10 Developing systems of support and reinforcement to those individuals undertaking evaluative efforts
16.11 Obtaining from select members of the educational milieu reactions to facilitative efforts
16.12 Facilitating, generally, the efforts of select members of the educational milieu to develop solutions for operating problems
16.13 Obtaining resources and resource personnel for the facilitation of evaluation efforts, the coordination of evaluation efforts and for decision-making efforts.

D. Evaluation of Students and Instruction

A third component of the evaluation conceptual paper with its resulting pool of objectives focuses on student and instructional evaluation. An experienced psychologist was asked to develop a set of instructional objectives for training personnel in the area of student evaluation. We hoped that any evaluation objectives missed by the Welch and Coller taxonomies would be picked up by this third effort. The results of that effort follow as Part D of the Objectives for Evaluation.

It should be pointed out that the rationale for the following report varies somewhat from previous efforts. A self-analysis has been combined with a search of the literature to develop a model for evaluation based on decision making. The purpose and nature of these decisions leads to a set of evaluation objectives. Although the path traveled is different than in the previous papers, the kind of terminal behavior anticipated is very similar to the other papers.

This part of the conceptual paper is presented as originally written by Terwilliger. It is followed by a set of six examples by which these objectives can be translated into terminal behaviors. The combined set of objectives, that is, those developed by Welch, Coller, and Terwilliger form the basis for the development of the training modules.

Introduction (Terwilliger, 1970)

There are many different decisions made about (and by) students in the typical school setting. Any choice which affects an individual's subsequent experience within the school (and, very likely, after the student leaves the school) is such a decision. As examples, a student may: (1) be assigned to a math section for "average ability" students, (2) choose a course in metal working as an elective, (3) be selected to participate in a special work-study program designed to augment the vocational training offered by the school, and (4) be assigned a grade of "C" by his English teacher.

Each of these decisions requires that a judgment (evaluation) be made. Ideally, such judgments are based upon a well-defined data base. The ultimate objective of all evaluation procedures is to arrive at decisions which are in some sense "better" than those decisions which would have otherwise been made.

Cronbach and Glaser (1965) have discussed ways for defining the "goodness" of a decision. They note that decisions are typically judged according to their benefit or utility in relation to possible alternatives. The benefit or utility of a decision may be defined with reference to a set of values possessed by the school or with reference to the values possessed by the individual student. Cronbach and Gleser refer to these two frames of reference as institutional decisions and individual decisions, respectively.

A major difference between institutional and individual decisions is the frequency with which they occur. From the institutional point of view decisions are repeated over and over so the decision maker usually searches for a decision strategy which will maximize

the benefit from a whole series of decisions. Since each decision involves the same set of values, different decisions can be combined to yield some overall outcome. The preferred decision rule is the one which will work best "on the average". Since institutional decisions can readily be formalized through mathematical and statistical techniques which make generalization possible they have become the major focus of psychometric theory.

Viewed from the perspective of the individual, a decision is often unique. A particular choice may occur only once. The "best" course of action depends upon the individual's value system and varies from one individual to the next. A particular objective or goal may be highly valued by one person but have little value to another. Because the individual decision occurs with such low frequency, it is not realistic to think of an "average" outcome. Consequently, systematic procedures for formulating and specifying individual decision values which are optimal are virtually non-existent at present.

Institutional decisions may be divided into three broad classes designated by Cronbach and Glaser as:

1. classification decisions
2. placement decisions

and

3. selection decisions

All three types of decisions involve choices among courses of actions or treatments (in the most general sense of the term). The classification decision is characterized by situations in which students are separated into several distinct treatment groups that are considered to be qualitatively distinct. Data used in classification decisions may come from traditional psychometric measures or may be

qualitative in nature. Typically, classification decisions are based upon multivariate data which combine both continuous variables and judgmental criteria. Examples of classification decisions are readily found in industrial or military settings where a common problem is assigning individuals to job specialities so as to maximize the benefit to the organization. However, in the school setting qualitative distinctions tend to rely upon individual student preferences (frequently tempered with guidance from a teacher, counselor, or parent) rather than institutional assignment. For example, a student may: (1) choose all possible courses in advanced mathematics in anticipation of college major in engineering, (2) elect subjects such as typing, shorthand, and office practice to prepare for a role as a secretary, (3) concentrate in the area of technical-vocational training insofar as the curriculum provides for this, or (4) select a broad range of subjects which will provide an appropriate background for college work in liberal arts.

Placement decisions represent a special type of classification problem. The placement decision is characterized by the fact that distinctions are made according to a single presumed underlying dimension. Even if the initial data are multivariate all the information on each individual is combined into a single composite index before a placement decision is made. A decision rule consists of determining how the univariate dimension is to be partioned for purposes of assigning individuals to treatments. It is important to note that the use of a single dimension implies that the differences among the treatments can be thought of in relation to presumed quantitative differences among the individuals.

Placement decisions abound in the educational context. At a general level, placement decisions include: (1) use of psychometric and other procedures to distinguish students (i.e., severly retarded) requiring special educational settings from those who do not, (2) assignment of students to specific groupings (i.e., accelerated, regular and remedial) for instructional purposes, and (3) individualization of instruction according to level of proficiency or aptitude demonstrated. Within a specific classroom setting, a teacher may make placement decisions by forming informal instructional groups according to judged ability level, prescribe certain review of remediation for individual students, or design a "contract system" which allows different students to work at different levels or paces.

A selection decision is defined by the existence of only two treatments, accept and reject. Like classficiation decisions selection decisions may be based upon multivariate information which combines both continuous and qualitative data. Although selection decisions are common in post school employment and in advanced educational programs (i.e., colleges, technical schools, and conservatories) they occur infrequently in public schools. Illustrations would be special courses or programs designed only for students who have been identified as especially talented or gifted in some specified field such as science, music, art, etc.

## Purposes for Decisions About Students

The motivation for classification, placement, and selection decisions is very different in the public school than in the military or industrial setting. Generally speaking, there are three major purposes for making decisions about students: (1) planning instruction, (2) monitoring instruction, and (3) assessing the result of instruction. These can be arranged in a cycle as shown in Figure 1.

The first stage of the cycle involves all decisions made in the course of planning instruction. This may consist of selecting students for specialized programs, placement of students by level of ability, or, in some cases, classification of students into broad groupings such as college preparatory, technical vocational, business vocational, etc. The second stage occurs during the instructional process when decisions are frequently made to alter or modify the instruction for specific students. This may take the form of a revision of pre-instructional decision, e.g., change the initial placement of a student. It may consist of prescribing remedial study or instruction for certain students. The monitoring decision may be to advance students through programmed stages in a planned instructional sequence. The third stage of the cycle is the assessment of terminal performance. This typically takes the form of constructing a composite or cumulative index of achievement which is translated into a letter grade or some other symbol that is communicated to parents and recorded on the student's permanent record.

There are two important features of Figure 1 that should be noted. First, as indicated by the figure, planning, monitoring, and assessing occur in a repeating cycle. The assessment of a student's performance at one point in time usually has some bearing upon the planning of his instruction at a later point. Therefore, the assessment of student performance is not a terminal decision but is one link in an endless chain of decisions. Second, the cycle depicted in Figure 1 may be considered on different time scales. It is common to equate the cycle with a school year in which students are assigned to class groupings in September, instruction occurs from September to June, and assessment occurs in June. It is just as easy to view the cycle on a more compact time scale. A teacher may complete the same cycle in teaching a unit which lasts only two or three weeks. Typically, the planning, monitoring, and assessment cycle is repeated several times on a reduced time scale within each larger time cycle. Also, on a reduced time scale the cycle may not involve all three stages, e.g., no differentiation of students may take place prior to instruction.

## Models for Making Decisions About Students

The rationale employed in making decisions can (and frequently does) vary from one situation to the next. Rationales or models for arriving at decisions about students typically reflect the edueational philosophy of the particular administrator, teacher, counselor, etc. making a decision, as well as the purpose for which the decision is being made. There are three general classes of models for decisions about students.

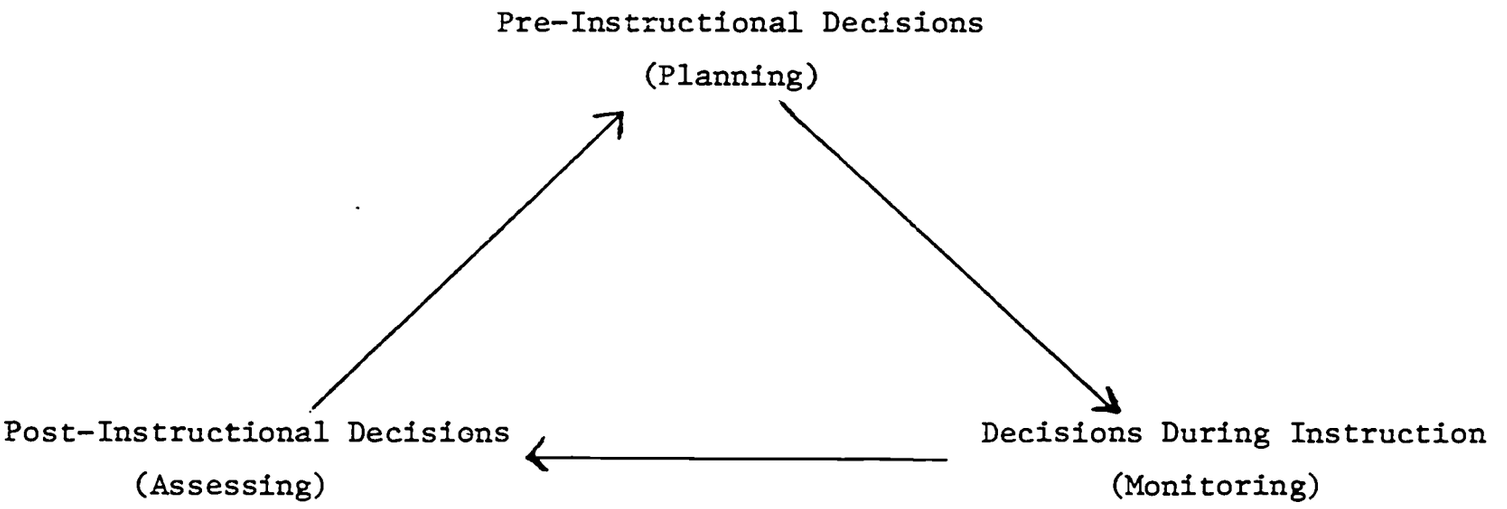



Figure 1 The Decision Cycle

These are:

1. criterion - referenced models
2. norm - referenced models

and 3. self - referenced models

Criterion - referenced models have received a great deal of attention recently. The primary feature of criterion - referenced models is that decisions are based upon a priori criteria in terms of observable behavior. Learning theorists and proponents of programmed instruction such as Gagne (1965), Glaser (1968), and Atkinson (1967) have stressed this approach. Bloom (1968) is probably the most vocal advocate of criterion - referenced models. He has described in some detail what he terms a model for "Learning for mastery". His notion of mastery is based upon the assumption that subject matter can be arranged into a meaningful hierarchy of complexity. Much of his thinking is based upon an earlier paper by Carroll (1963) which presents a model for school learning that differs in significant ways from traditional models of learning. Hively, et. al. (1968) have also described a criterion - referenced model which is based upon the concept of sampling from a well-defined domain of content.

Norm - referenced models are largely derived from psychological theories regarding the nature of individual differences. Most large - scale aptitude and achievement testing programs presently employed in schools reflect such thinking as does traditional measurement theory in psychology and education. The distinguishing feature of the norm - referenced model is that decisions are based upon the performance of a student in relation to some specified reference group.

Self - referenced models represent an attempt to individualize the basis for decisions. There are two major variations among this class of models. First, the notion of growth or improvement has been frequently suggested as the proper basis for making many decisions about individual students. Second, the idea of relating observed performance to some expectation based upon previous knowledge of the student (e.g., his general ability or a specific aptitude score) has also received some support. Self - referenced models have been especially popular among those who are concerned with diagnosing educational problems and prescribing programs for remediation.

## Training Educational Evaluators to Make Decisions About Students

When decisions about students are considered with reference to type, purpose, and model, the three dimensional solid shown in Figure 2 emerges. There are 36 (4x3x3) cells in the figure, each one representing a unique combination of type, purpose, and model. It is obvious that certain cells are represented more frequently in decisions about students than are others. Indeed, it is difficult to imagine instances of certain combinations, e.g., classification in assessing performance according to a self - referenced model. However, the three-way conception is helpful in pointing out the many forms which decisions take.

The prospective educational evaluator must learn to identify specific decision problems according to the schema shown in Figure 2. Further, he must recognize the implications for activities in which he will be engaged.

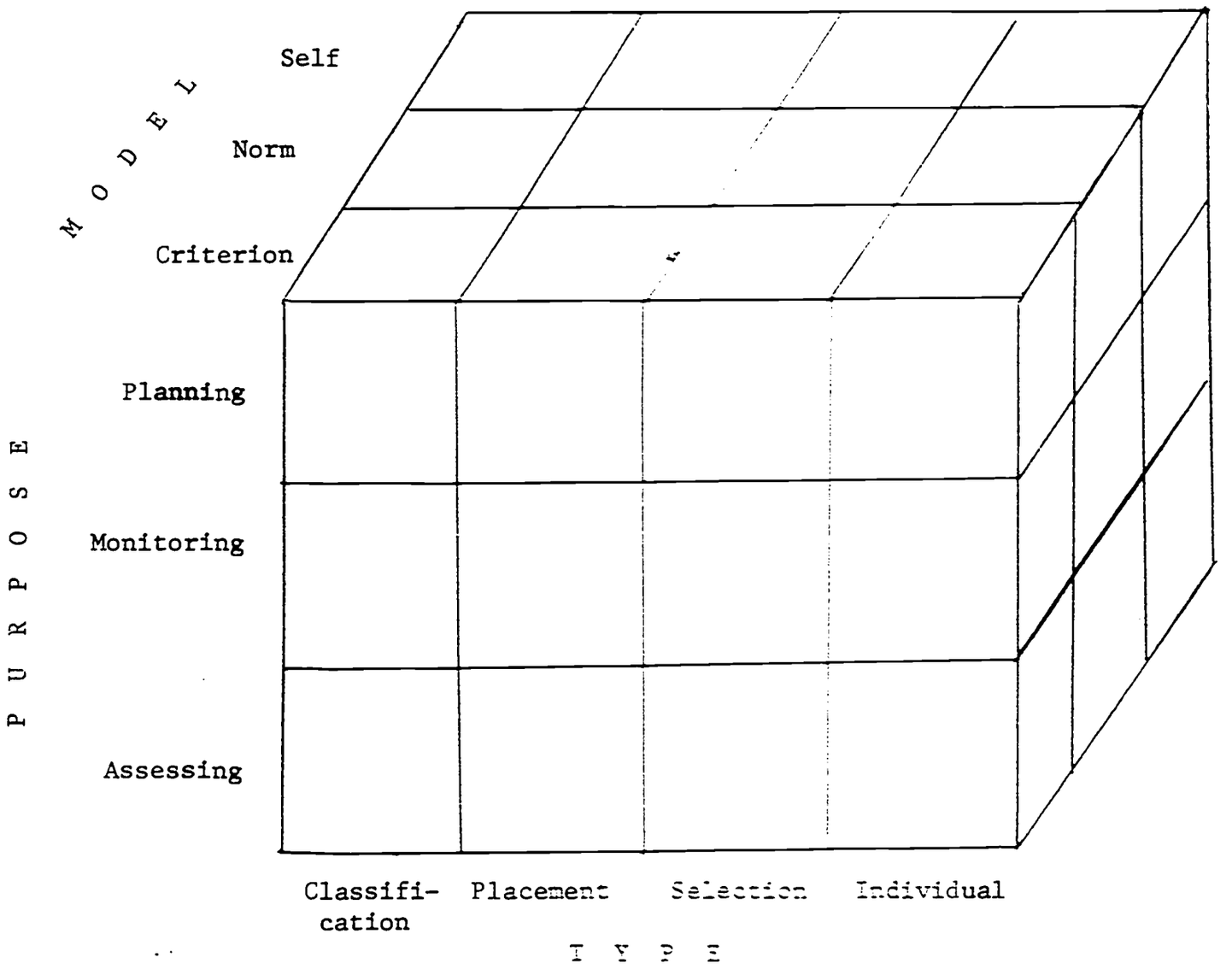



Figure 2 Decisions Classified by Type, Purpose, and Model

Differences among the possible kinds of decisions represented in the figure have a direct bearing upon: (1) the nature and amount of data to be collected, (2) the way in which the data are treated, and (3) the criteria by which decision-making strategies are judged. The general schema of Figure 2 is related to a hierarchy of general evaluation skills in Table 1. The hierarchy is the same one previously discussed in reference to program evaluation.

Note: The set of objectives found in Table 4 was developed independently of the Welch hierarchy, yet nearly all objectives fall at one of the 12 levels. For the convenience of the reader, the objectives are keyed to the hierarchy.

It should be pointed out that the component skills for each level are not exhaustive, but rather indicative of the content of the training modules. Further work is required once the training concept is accepted.

Table 4

| Objectives Keyed to Welch Hierarchy Levels | | |
|---|---|---|
| | A. | Requisite Skills for Planning an Evaluation Project |
| II. | 1. | defines different types, purposes, and models for decisions |
| II. | 2. | given a specific setting, correctly identifies type and purpose of the decisions |
| II. | 3. | given a specific setting, determines appropriate model to employ |
| II. | 4. | given a specific setting, determines appropriate data to collect |
| II. | 5. | designs one or more decision strategies for a given decision problem |
| II. | 6. | proposes specific ways to evaluate and/or modify a decision strategy |
| | | a. employs concepts of cost, payoff, and expected utility |
| | | b. relates traditional psychometric concepts (e.g., reliability and validity) to decision strategies |
| -- | 7. | designs original techniques for evaluating decision strategies |
| | B. | Requisite Skills for Data Collection |
| XI. | 1. | administers standardized group tests |
| XI. | 2. | administers standardized individual tests |
| XI. | 3. | conducts interviews |
| XI. | 4. | employs observational (rating) techniques |
| VIII. | 5. | appropriately applies concepts basic to measurement (e.g., reliability and validity) |
| IX. | 6. | makes appropriate selection of standardized measures of achievement, aptitude, and interest to solve specific problems |
| VIII. | 7. | constructs specific measures of achievement, or attitudes |
| | | a. designs instruments using criterion-referenced model |
| | | b. designs instruments using norm-referenced model |
| | | c. designs instruments using self-referenced model |
| VIII. | 8. | constructs non-test data collection instruments (e.g., designs rater observation techniques) |
| | 9. | trains others in use of data collection techniques |

Table 4 (cont'd)

| Objectives Keyed to Welch Hierarchy Levels | | |
|---|---|---|
| | C. | Requisite Skills for Data Preparation |
| XII. | 1. | scores responses to choice type items using key |
| XII. | 2. | scores free response protocal using guide |
| XII. | 3. | designs card layouts and punches raw data and format statements |
| XII. | 4. | prepares response protocals for machine scoring |
| -- | 5. | writes computer program to perform scoring |
| XII. | 6. | devises scoring procedures for complex performance which are judged or rated |
| -- | 7. | trains others to perform scoring operations |
| | D. | Requisite Skills for Data Analysis |
| VII. | 1. | defines basic descriptive statistics commonly employed with score distributions (e.g., mean, std. dev., median, Q) |
| VII. | 2. | given specific data, constructs histogram or frequency polygon |
| VII. | 3. | given data (or a graph), correctly describes distribution in terms of shape, central tendency, and variability |
| VII. | 4. | given specific data, calculates (by hand or desk calculator) mean, std. dev., median, and Q |
| VII. | 5. | distinquishes between correlation (r) and regression problems and gives an example of each |
| VII. | 6. | given specific data, calculates the correlation (r) between two variables and determines the regression equation in predicting one from the other |
| VII. | 7. | performs descriptive statistical analyses (univariate and bivariate) with the aid of standard computer programs |
| VII. | 8. | defines and cites examples using basic inferential concepts (e.g., standard error, null hypothesis, etc.) |
| VII. | 9. | given data, constructs contingency (or expendency) tables and describes these in terms of probability statements |
| VII. | 10. | given data, calculates more commonly used inferential statistical tests (e.g., t-test, simple, ANOVA, and certain non-parametric tests |
| VII. | 11. | performs inferential statistical analyses with aid of standard computer programs (objective C-3 is prerequisite) |
| VII. | 12. | derives special statistical indices for unique problems |

Table 4 (cont'd)

| Objectives Keyed to Welch Hierarchy Levels | |
|---|---|
| | E. Requisite Skills for Data Interpretation |
| IV. | 1. distinguishes between descriptive and inferential interpretation of data |
| IV. | 2. given adequate information, provides plausible explanation for results (descriptive statistics only) |
| IV. | 3. given adequate information, provides plausible explanation for results (inferential statistics) |
| IV. | 4. given inferential analyses, draws conclusion consistent with the data |
| IV. | 5. recognizes faulty inferences based upon data presented |
| | F. Requisite Skills in Reporting Projects |
| VI. | 1. describes accurately and concisely completed or planned projects |
| VI. | 2. relates project to other relevant studies |
| VI. | 3. draws appropriate conclusions from previous and current research |
| VI. | 4. proposes specific courses of action to decision makers on basis of research findings |

E. Translating Instructional Objectives Into Terminal Behaviors

In each of the preceding sets of evaluation objectives, functions, tasks, or whatever term is chosen we see considerable overlap. This is appropriate. However, there are differences among the schemes and resolving these differences will fall heir to the implementation stage. However, these schemes, together with the additional work to grow out of the Schalock study, the AERA papers, and the UMC consortium, makes it quite clear that a definitive set of evaluation objectives will be created.

Following agreement upon the tasks of the evaluator will come the development of specific training modules to accomplish the desired terminal behaviors. Considerable research on this phase will occupy much of the early effort of the implementation stage. The consortium has excellent facilities and personnel for conducting this research to assure that a prescribed training package adequately results in the expected terminal behavior. The plan for accomplishing this consists of specifying the objective, the entry skill conditions, expected terminal behavior, a measure for observing this behavior, and suggested materials for training. Following are listed examples for several objectives selected from the preceeding hierarches. These should give the flavor of the technique. It is expected that conditions, identifiable behaviors, and appropriate training modules can be developed to include each of the objectives defined for evaluation.

Objective A-2: (Level II) The trainee will correctly identify the type and purpose of decisions concerning students when provided with a "case description."

Conditions: Five brief case descriptions (one page) of specific school settings calling for decisions about students will be given. Trainees will be given ample time (one hour) to read the descriptions and make responses.

Behavior: For each case, the trainee is to specify in writing the type of decision (classification, placement, selection) and the purpose of the decision (planning, monitoring, assessing).

Measure: Two points will be awarded for each correct designation of a case (one for type and one for purpose). A minimal acceptable score is eight points (out of a possible ten).

Materials: A written unit on types and purpose of decisions about students will be used for training. This unit will provide several examples appropriate to the school setting.

Objective A-5: (Level II) The trainee will design one or more decision strategies for a given decision problem.

Conditions: A general decision problem requiring assignment of students to different treatments will be described in writing (one or two pages). Trainees will be given sufficient time to carefully read the materials and structure a response, e.g., 2-4 hours.

Behavior: (Product) Trainees are to outline in writing (and schematically) the specific procedures they would follow in designing a decision strategy for assigning students to treatments. Specifically, the response should consider: (1) the input variable(s), (2) the outcome variable(s), (3) the way in which input variables are to be translated into decision making, and (4) the approximate cost of data collection and analysis.

Measure: The response of each trainee will be rated independently by two qualified trainers with respect to specific criteria. Each trainer will also make a global Satisfactory-Unsatisfactory judgment. If both judges agree on the global judgment, the response will be evaluated consistent with their evaluation. If the two judges disagree, they will resolve their differences through a joint review of the response.

Materials: Specially prepared units adapted from sources such as Cronbach and Gleser. Readings and lectures in E Psy 117 or equivalent introductory course on principles of measurement.

Objective B-7: (Level VIII) The trainee will construct a specific measure of achievement (or attitude) according to acceptable psychometric principles.

Conditions: Each trainee will choose as a project the design and construction of a paper and pencil measure of achievement or attitude for use with students. The task will be a "take home" project embedded within a formal course on instrument construction. The project will take a period of 4 to 6 weeks.

Behavior: (Product) The trainee will produce a statement that explains: (1) the purpose of the instrument and the setting in which it is to be used, (2) the general logic or model (criterion-referenced, norm-referenced, or self referenced) assumed in the construction of the instrument, (3) the way in which performance on the instrument can be related to a decision about students, and (4) proposed methods for determining whether the instrument functions as intended (e.g., validation studies). A copy of the instrument, directions for administration, and procedures for scoring will also be required.

Measure: The instrument and associated statement will be judged with respect to clarity, logic, and overall quality by the course instructor. A judgment of acceptable vs. unacceptable will be assigned to the total project.

Materials: Selected references on instrument design and construction will be required reading. In addition, formal training sessions (lectures and demonstrations) dealing with instrument construction will be employed in a course devoted to construction of educational measures.

Objective C-2: (Level VII) The trainee will score a free response test protocal using a written scoring guide.

Conditions: Three written response protocals and the scoring directions for a commercially available free response measure (e.g., Torrance Tests of Creative Thinking) will be given to the trainee. The recommended time will be allowed for scoring each protocal.

Behavior: The trainee is to score each protocal in accord with the directions provided and determine the composite score using the scoring formula given in the manual.

Measure: The scoring on each protocal will be compared with that of an experienced scorer. The "percent of agreement" between the experienced scorer and the trainee will be calculated for each protocal. A minimal acceptable performance is 90% agreement on each of the three protocals.

Materials: A standard manual and scoring guide for the instrument employed will be provided for each trainee. At least five trial protocals will be scored prior to the three employed for testing. Special training sessions will also be held to discuss particular scoring problems.

Objective E-4: (Level IV) Given results from a statistical analysis, the trainee will state conclusions consistent with the data (Analysis of contingency table by $\chi^2$ used as illustration).

Conditions: A detailed description of a particular educational problem involving an analysis of a contingency table will be given. The contingency table, $\chi^2$ computations and $\alpha$ level adapted for the test will also be shown.

Behavior: The trainee is to determine the appropriate number of degrees of freedom and state whether the observed $\chi^2$ value is significant at the stated $\alpha$ level. A statement of conclusions relevant to the problem is initially posed will also be required.

Measure: The trainee's statement of the number of degrees of freedom, significance of the observed $\chi^2$ value, and conclusions concerning the problem will be checked for accuracy. A minimal performance is the correct value of the degrees of freedom and appropriate decision regarding rejection of the hypotheses at the stated $\alpha$ level.

Materials: The $\chi^2$ distribution, its interpretation and problems involving applications to contingency tables will be presented in a course on statistical inference, e.g., E. Psy. 217.

Objective D-11: (Level VII) The trainee will perform inferential statistical analysis with the aid of standard computer programs. (Simple ANOVA used as illustration.)

Conditions: Raw score distributions for four different groups (unequal frequencies) will be presented with instructions to perform a simple ANOVA. A manual of computer programs with accompanying instructions for their use will be provided.

Behavior: The trainee must do the following:
(1) design a data card layout and punch raw data cards
(2) write and punch appropriate format statements
(3) punch necessary control card(s) in accordance with the program manual
(4) punch necessary system control cards
(5) submit complete job deck to the computer and pick up output for inspection

Measure: A checklist procedure will be used to determine errors (if any) at each step in the sequence outlined above. If errors occur, the trainee must resubmit the job until correct output is obtained.

Materials: The necessary conceptual materials for understanding the statistical analysis will be presented in a course in statistical inference e.g., E. Psy. 217. The skills in card preparation and punching will be acquired with the aid of programmed materials and practice problems.

F. Bibliography

Bloom, B. S., "Learning for Mastery", Evaluation Comment, Center for the Study of Evaluation, University of California at Los Angeles, 1968, 1, No. 2.

Carroll, J. B., "A Model of School Learning", Teachers College Record, 1963, 64, 723-733.

Coller, Alan, "A Taxonomy of Programmatic Tasks in an Educational Evaluation Facilitation and Coordination System", CIRCE, School of Education, University of Illinois. (Mimeo) August, 1969, (Revised July, 1970).

Cronbach, L. J. and G. C. Gleser, Psychological Tests and Personnel Decisions. (2nd Edition) University of Illinois Press, 1965.

Gagne, R. The Conditions of Learning, Holt, Rinehart, Winston, 1965, 308 p.

Glaser, R. "Adapting the Elementary School Curriculum to Individual Performance", Proceedings of the 1967 Instructional Conference on Testing Problems. Educational Testing Service, 1968.

Hively, W., et. al. "A 'Universe-defined' System of Arithmetic Achievement Tests", Journal of Educational Measurement, Winter, 1968, 5, 275-290.

Schalock, H. Del, et. al., "Interview Schedule B: Form III-R -- Evaluation", RDD & E Manpower Study Project. 1970 (Mimeo).

Welch, Wayne W., "Conceptual Framework -- Training Educational Evaluators", College of Education, University of Minnesota, Minneapolis. (Mimeo) August 1970.

Welch, Wayne W., "A Summative Evaluation of the PSNS Course. Part I, Strategy and Implementation", (Mimeo), College of Education, University of Minnesota, Minneapolis, 1970.

Design Document II
Appendix I

# A Framework for Writing Behavioral Objectives for Skills and Knowledges of Educational Research

Paul E. Johnson
University of Minnesota
Consultant for Project 70-12
Upper Midwest Regional Educational Laboratory

Note:

It is again brought to the readers attention that a decision was reached part way through the design phase not to focus our training efforts on research per se. However, the following paper had been written prior to this decision. Because of its rather interesting approach to the problem and the value of the listed terminal behaviors for the evaluation effort, it is included as an Appendix to Design Document II.

Many of the objectives derived here will be included in the training modules developed for evaluation training. Furthermore, the DAT coding system employed by Johnson may have applicability for the total project. Meanwhile this approach to defining the domain of skills for educational researchers is presented as an illustration of our attempt to design "new patterns of training."

Index

## The Framework

The successful training program is usually one with some rationale for selecting performances which are to be incorporated into training procedures. The program proposed here uses the idea of ability or competence to facilitate the generation of training tasks (performances) as well as the computer storage of these tasks.

Competence for our purposes is a way of talking about maximum, potential or idealized behavior. It is something which is seldom, if ever, realized in practice. Because competence represents idealized behavior it lends itself to formal definition; and it is in formal definition that we find the tools for generating performance.

The training program we propose is designed to produce behavioral engineers and technicians, not behavioral scientists. Thus, the domains of performance we wish to achieve do not represent any single subject matter, rather, we conceive of them as subject matter fragments. Three categories were constructed to describe the subject matter fragments in our program. These categories are labeled design, analysis, and theory.

Design gives us the means for arranging conditions to accomplish some specified goal or task. Design includes rules for determining cause and effect relationships as well as schemes for organizing information into categories in order to achieve economy in description and communication.

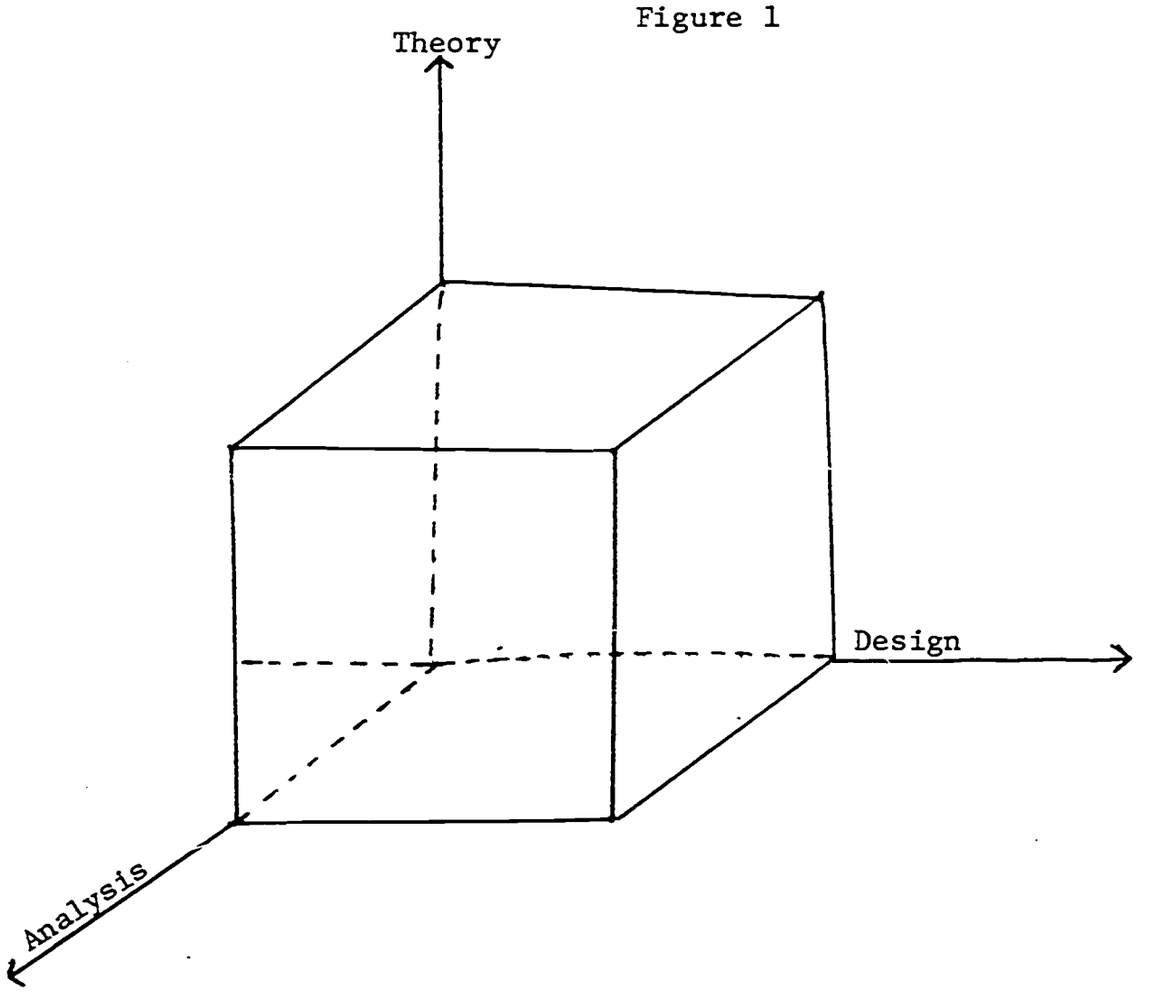
Theory gives us a means for selecting conditions and tasks in relation to the goals to be achieved. In a formal sense, theory

involves the use of concepts to accomplish description and explanation. Less formally, theory is a means of accounting for the outcome of tasks.

Analysis gives us tools for examining the results of performance to see whether goals have been achieved. These tools include procedures for assigning numbers so that we can make an empirical assessment of success or failure as well as logical schemes for dividing tasks into meaningful parts.

We conceptualize the relationship among the three categories as shown below.

Figure 1




According to this figure a particular competence (what we shall call a task structure) is specified as an intersection of one or more of the three categories.

Given that our training problem is one which requires that individuals (trainees) be competent in some of the skills of educational research, we can define the three categories as follows.

Design

Level 0 Representation of Relationships (Graphs, Tables, Charts)

Level 1 Functional Relations

Independent variables, dependent variables

Level 2 Sampling

Random, Stratified

Level 3 Logic of Design

Experimental variables, control variables, confounding variables

Level 4 Designs

Two groups, multiple groups factorial

Analysis

Level 0 Representation of Relationships

Level 1 Organization of Information

Logic of Categorization, labeling

Level 2 Measures

Level 3 Distribution

Level 4 Measures of Distribution

Central tendency, variability

Level 5 Determining Cause and Effect

Testing hypotheses

Level 6 Measures of relationship

Theory

Level 0 Representation of Relationships

Level 1 Operational Definitions

Level 2 Description

Level 3 Prediction

Level 4 Prescription

Level 5 Explanation

All three dimensions in the figure on page 2 are formulated as ordinal scales with the basic or fundamental competencies at the lowest level. If the three dimensions are to have a common origin they must have a competence in common. For purposes of our example this competence is the ability to construct graphs, tables, and charts in order to represent the covariation of one variable with another, or the way one variable changes with time.

We use the figure to generate task structures which specify the competence underlying specific training performances. This is accomplished by assigning each point in the space a DAT Code which represents its position on the three dimensions. These positions are recorded as follows:

| DAT Code | Task Structure |
|---|---|
| 4, 6, 5 | Designs, Measures of Relationship, Explanation |
| 4, 6, 4 | Designs, Measures of Relationship, Prescription |
| 4, 6, 3 | Designs, Measures of Relationship, Prediction |
| 4, 6, 2 | Designs, Measures of Relationship, Description |
| 4, 6, 1 | Designs, Measures of Relationship, Operational Definition |
| 4, 6, 0 | Designs, Measures of Relationship, Representation of Relationships |

| | |
|---|---|
| 4, 5, 5 | Designs, Cause and Effect, Explanation |
| 4, 4, 5 | Designs, Measures of Distribution, Explanation |
| 4, 3, 5 | Designs, Distributions, Explanation |
| 4, 2, 5 | Designs, Measures, Explanation |
| 4, 1, 5 | Designs, Organization of Information, Explanation |
| 4, 0, 5 | Designs, Representation of Relationships, Explanation |
| 3, 6, 5 | Logic of Design, Measures of Relationship, Explanation |
| 2, 6, 5 | Sampling, Measures of Relationship, Explanation |
| 1, 6, 5 | Functional Relations, Measures of Relationship, Explanation |
| 0, 6, 5 | Representation of Relationships, Measures of Relationship, Explanation |
| . . . | |
| . . . | |
| . . . | |
| 1, 1, 1 | Functional Relations, Organization of Information, Operational Definition |
| 1, 1, 0 | Functional Relations, Organization of Information, Representation of Relationships |
| 1, 0, 1 | Functional Relations, Representation of Relationships, Operational Definitions |
| 0, 1, 1 | Representation of Relationships, Organization of Information, Operational Definition |

It should be pointed out that the choice of categories to organize the subject matter fragments underlying training performance is not unique. These categories must be arrived at by requesting a sample of individuals (informants), competent in the skills and knowledge of research, to provide the basic or funda-

mental concepts around which knowledge and skills in research can be organized. By using the data from several individuals (a sample of informants) we arrive at a best or representative set of dimensions. Once this is accomplished we go back to our informants and ask them to list the competencies which make up each category. Finally, we ask our informants to order these competencies according to prerequisite knowledge or skill in order to obtain our ordinal scales.

While the categories represented in the preceding analysis are in a sense arbitrary, we view them as a first approximation to the basic or fundamental set of competencies which define the knowledge and skill required of personnel in research, evaluation and development. Furthermore, even though the categories were constructed to describe skills and knowledge in research, they seem to partially identify the competence underlying evaluation and development. In the case of evaluation, for example, one would suppose that a high priority is given to the dimension of analysis with less weight attached to design and least weight attached to theory. For development, however, theory may be most important, while design and analysis are of minimal importance.

The point is that in each case the dimensions which define competence must be specified. As more and more dimensions are defined in this fashion, however, we expect the basic set of competencies underlying training performance to become complete.

The proceeding task structures can be converted into performance by applying the rule "What should a person be able to do who knows X" where X is the task structure for a given DAT Code. For example, the DAT Code 4, 6, 5 would result in the statement "What should a person be able to do who knows: design, measures of relationship and explanation."

In order to generate training tasks we give the performance rule for each DAT Code, (and therefore each task structure) to a sample of personnel in educational research. We ask each individual in this sample to produce as many performances as he can. By taking the behaviors produced by a variety of individuals we are able to construct a more or less complete set of tasks for each task structure.

For anyone to generate tasks using the above rule we must specify characteristics of the desired behaviors. We assume that these characteristics (responses) are of three kinds: recognizing or identifying, selecting or classifying, and producing. Once the decision is made as to which form task performance is to take, actual behaviors can be specified.

To complete our description of tasks, we must specify the relevant stimulus events (conditions) under which performance is to occur. We classify these conditions into three categories: symbols, iconics (pictures) and objects.

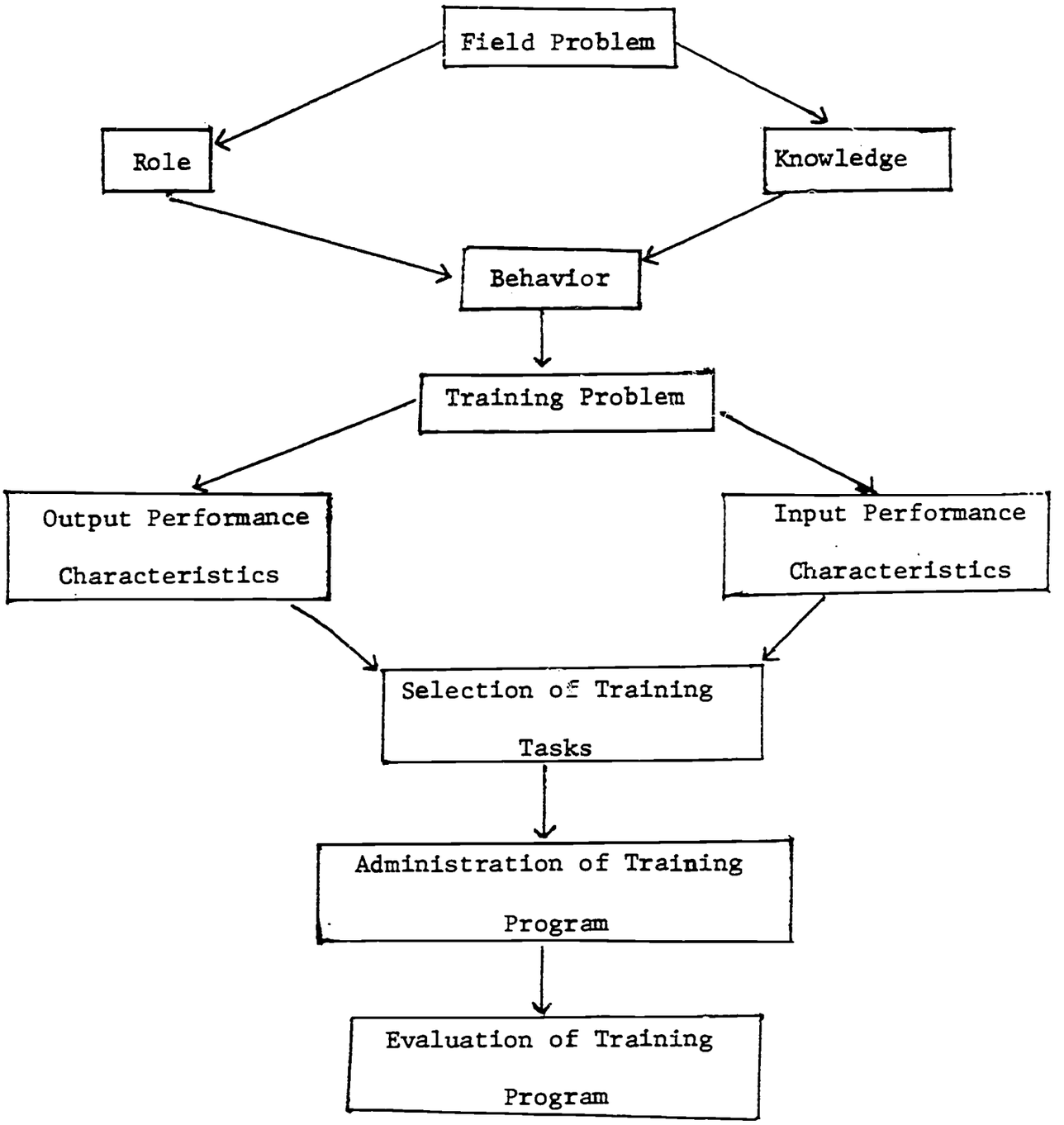
We represent the procedure for generating tasks as follows:

| | | Conditions | | |
|---|---|---|---|---|
| | | Symbols | Iconics | Objects |
| Response | Recognition | | | |
| | Selection | | | |
| | Production | | | |

There are nine different ways to represent a given task structure (and its corresponding DAT Code) in tasks (provided we do not have combinations of conditions or responses).

In practice we suppose that training problems arise from field problems through a specification of either roles or the desired knowledge or ability. In either case behaviors are generated which can be used to select a relevant training problem. Once the training problem is chosen we must specify values for the dimensions of design, analysis and theory which describe the knowledge and skill (competence) to be demonstrated. Next, we can construct the task structures which underly this skill or knowledge. Then we use the DAT Code for these task structures as prompts or cues in the performance rule "What should a person be able to do who knows X". And finally we produce actual tasks by selecting stimulus events and performance characteristics.

The following diagram is designed to illustrate how the framework developed here fits into the overall structure of the proposed training program.


Field Problem
Role
Knowledge
Behavior
Training Problem
Output Performance Characteristics
Input Performance Characteristics
Selection of Training Tasks
Administration of Training Program
Evaluation of Training Program

While the procedure for generating tasks for training and evaluation are relatively straightforward there are some points which need clarification. For example, since each point in Figure 1 on page 2 is an intersection of values on three dimensions, we should be able to generate performance for a given value on one dimension without knowing anything about values on the other dimensions. Thus, a task structure of design, measures of relationship, and explanation is equivalent to the three task structures: design, _____, _____; _____, measures of relationship, _____; and _____, _____, explanation. However, it seems reasonable to suppose that in some cases there may be an interaction among dimensions so that the competence required to know design, depends upon the level of competence on the analysis dimension.

To take account of such interactions we must present informants with combinations of task structure for any given DAT Code. Thus, an informant would be asked to generate performance for pure D, pure A, and pure T competence and also DA, DT, and AT combinations of competence as well as DAT competence.

If we suppose that ultimately the tasks produced by application of our framework are stored and retrieved by computer, then the computer address of a given task must indicate whether interactions between dimensions have been assumed. For example, the number 4, 0, 0 gives design tasks. The number 4, 6, 0 gives tasks which are assumed to require a knowledge of design and measures of relationship, while the address 4, 6, 5 gives tasks which assume a knowledge of design, measures of relationship and explanation. If we wish tasks for design, measures of relation-

ship, and explanation in a pure form we address them as 7, 0, 0; 0, 6, 0; and 0, 0, 5 respectively.

The next point that needs clarification is hierarchical dependencies among tasks. Ordinarily, one assumes that some performances are prerequisite for others, and this is true in our scheme as well. What we have added, however, is the idea that hierarchies also exist in the underlying competence. Thus, we assume that in order to know the logic of design, an individual must possess the knowledge below this competence on the design dimension. The behaviors generated to represent logic of design therefore assumes the behaviors which represent the competencies below it.

In addition to hierarchies of competence we also allow for hierarchical dependencies among performance within a given competence level. Here we assume that some performances which represent logic of design are prerequisite to other performances which also represent logic of design. The only way we have of determining such performance dependencies is by again using the judgment of informants. Our procedure is to present a given task to an informant and ask what performances are necessary or prerequisite to it. We intend to use tree structures to represent the relations obtained in this fashion. The computer address of a given task therefore includes a code which allows prerequisite performances to be obtained (see the matrix procedure developed elsewhere in the proposal for accomplishing this).

The above procedures for generating tasks from a specification of an underlying competence do not apply directly to the

training of para professional personnel. In this case we believe the desired competence is better thought of in terms of specific task requirements. For example, to determine the training performances for key punch operators or test administrators we simply present the verbal label for the task to informants, together with a performance rule which includes conditions and responses. Once we have a more or less complete set of tasks, hierarchical dependencies can be generated as before by asking informants to supply prerequisite behaviors where they have not been provided and order those which are already given.

## Defining Objectives

A behavioral objective consists of a specification of behaviors and the general conditions under which they are to occur. In our framework behavioral objectives are represented by the nine cells in the table on page eight.

Generating tasks means operationally defining the conditions and responses on the margins in the table. For example, we could decide that by iconics, we mean graphs; by symbols, we mean words; and by objects, we mean standardized tests of subject matter achievement. We could further decide that by recognition responses, we mean identifying a statement as a member of a category (e.g., correct or incorrect); by selection responses, we mean choosing from among a specified set of alternatives as in a multiple choice test; and by production responses, we mean drawing graphs.

The following two examples are designed to illustrate the

development of behavioral objectives and tasks from our framework.

Case 1

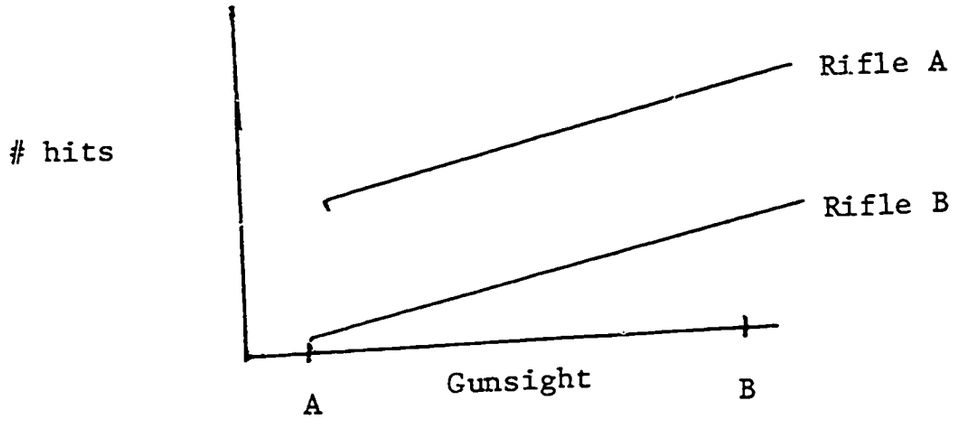
Suppose we wish to produce tasks which represent logic of design, representation of relationships, and description (DAT Code 3, 0, 2). Suppose further we decide to employ iconic conditions and a recognition response. The performance rule might then produce the following.

Present subjects (trainees) with the statement.

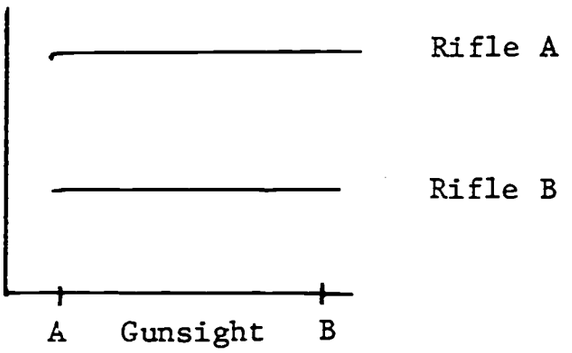
"We are given two types of rifles and two types of gunsights. Each rifle is fired with each gunsight five times. The number of hits on a target for each of the four conbinations is recorded on a graph."

Next, present subjects with the following graphs one at a time and ask them to identify those graphs which indicate an interaction between type of rifle and type of gunsight.

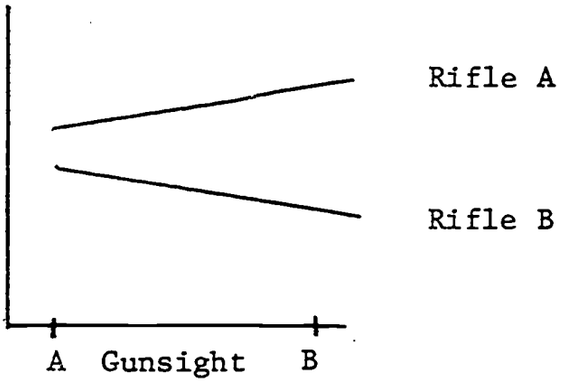
Graph 1



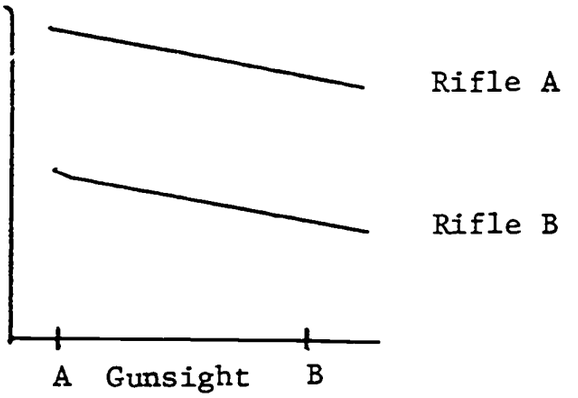
Graph 2



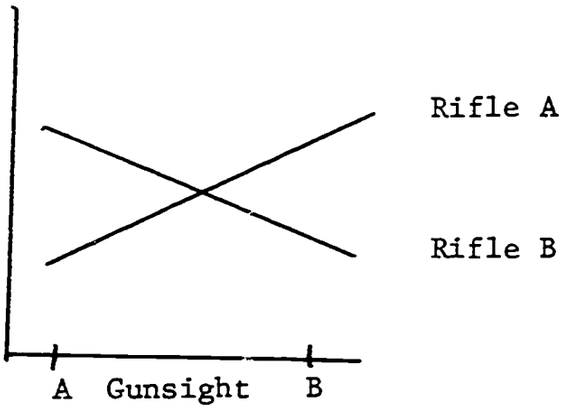

Graph 3




Graph 4




Graph 5



We could change the task slightly to one involving a production response by requesting subject to draw several graphs which show an interaction between types of rifle and type of gunsight.

Case 2

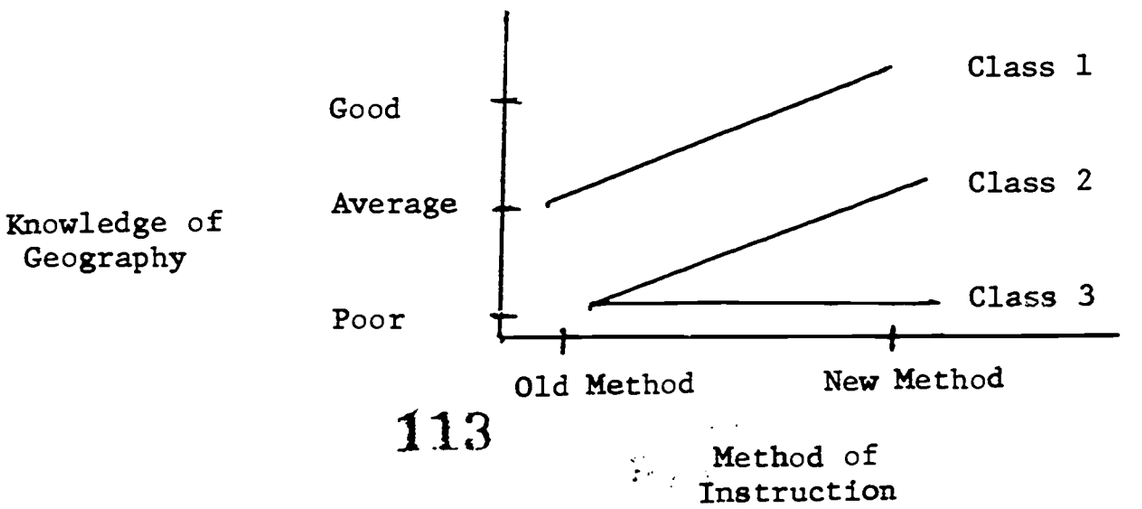
Suppose we wish to write tasks to represent the DAT Code 1, 1, 2 which has the task structure: functional relations, representation of relationships, and description. Suppose further that we wish to use iconic conditions and a selection response. The following training item might then be produced.

Present subject (trainees) with the statement:

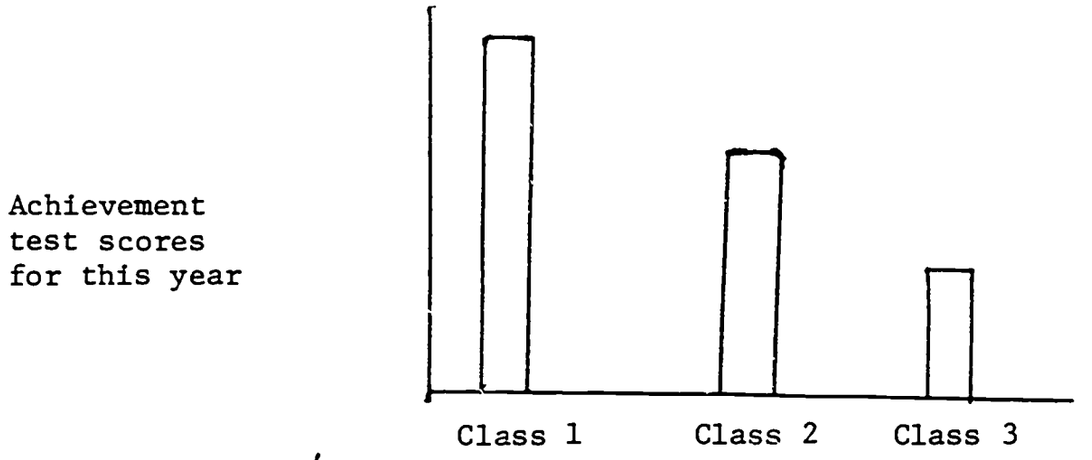
"We administer a standardized achievement test in geography to students in several geography classes in a suburban high school. We wish to present the results of this test to the board of education to demonstrate that a new method of instruction employed in che geography classes has resulted in students having a greater knowledge of geography."

Next present subjects with the following graphs and ask them to select the one that they should present to the board of education.

Graph 1

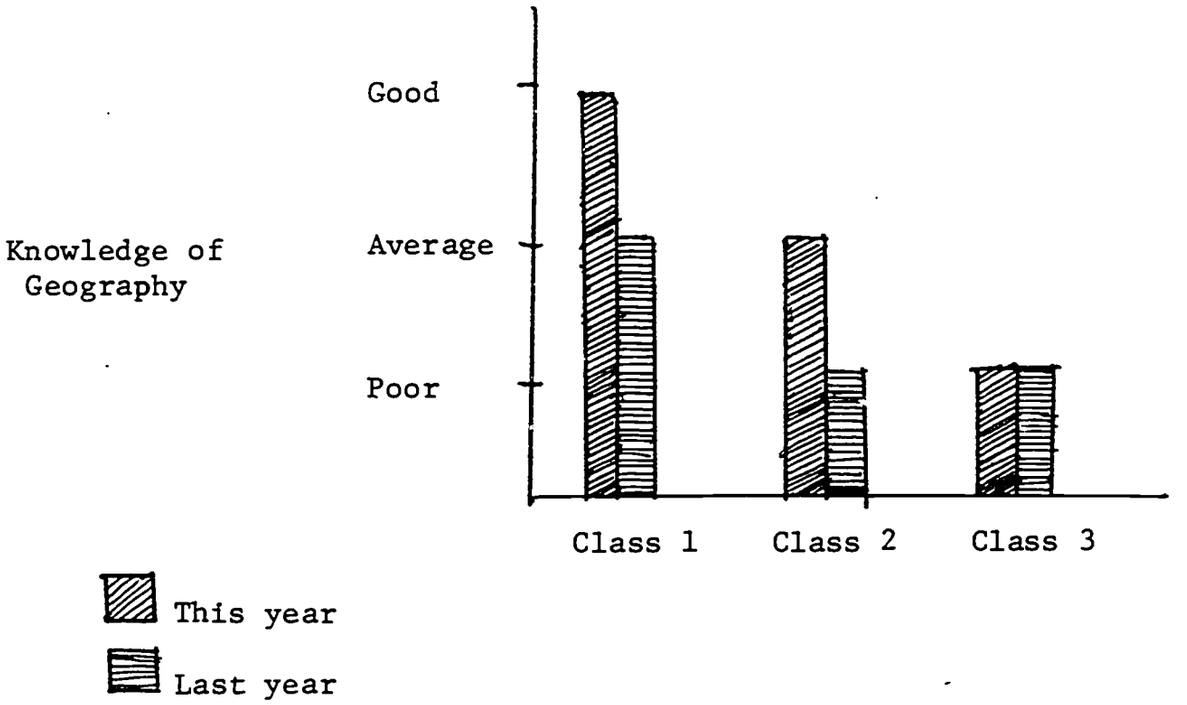


Graph 2




Graph 3


Graph 4




As in Case 1, we could convert Case 2 to a production task by requiring that the subject draw the desired graph. We could also make

Case 2 a recognition task by simply giving subjects the four graphs one at a time and asking them to identify those that represent the information needed for the board of education.

The above examples are intended to be suggestive rather than exhaustive. In both cases many instances such as those presented could be constructed. The important thing to realize is that our procedures are not intended to be radical departures from the usual methods of arriving at tasks for purposes of teaching and evaluation. Rather, we view them as an attempt to make explicit the intuitive ingredients of writing behavioral objectives. We believe this is the only way the goals of training can be specified clearly and unambiguously enough to be achieved.

## Sample Objectives and Tasks

As described elsewhere in this proposal, we construct behavioral objectives by working downward in a hierarchy of definitions from the general to the specific. At the highest level of generality we have roles defined by field problems as illustrated on p. 9 of this section. Roles lead to conceptual objectives which in turn lead to educational objectives. These lead to instructional and behavioral objectives which lead finally to actual tasks.

To illustrate how this hierarchy can be applied we have developed a number of educational objectives. We state these without a specification of conditions or behaviors as follows.

Educational Objectives

| DAT Code | | |
|---|---|---|
| 302 | 1) | To design experiments which assess relationships between independent and dependent variables. |
| 101 | 2) | To recognize and define independent variables. |
| 101 | 3) | To recognize and define dependent variables. |
| 101 | 4) | To recognize and define a functional relationship. |
| 315 | 5) | To state and recognize the nature of the scientific process. |
| 415 | 6) | To integrate philosophical concepts and logical principles in the logic of experimental design. |
| 415 | 7) | To recognize the strengths and limitations of the scientific method as applied to experimentation. |
| 415 | 8) | To recognize and state the importance of controls in scientific experimentation. |
| 415 | 9) | To recognize and state the criteria for designing experiments. |
| 010 | 10) | To employ techniques of graphing |
| 041 | 11) | To recognize and state definitions for the terms graphing, score intervals, midpoints, and actual class limits. |
| 041 | 12) | To recognize and state parameters of distribution. |
| 041 | 13) | To recognize and state parametric assumptions of normality required by sample data. |
| 041 | 14) | To compute and interpret means, medians, modes. |
| 041 | 15) | To compute and interpret standard deviations and variances. |
| 041 | 16) | To compute and interpret standard scores. |
| 254 | 17) | To distinguish between population and sample characteristics. |
| 254 | 18) | To recognize and state relationships between z and t. |

254 19) To state and recognize the meaning of the null hypothesis.

315 20) To assess the number of degrees of freedom in a system.

041 21) To compute and interpret sums of squares.

051 22) To compute and interpret Pearson product moment correlation coefficients.

041 23) To compute and interpret percentile ranks, and standard scores.

051 24) To distinguish between correlation and regression.

525 25) To design an experiment to determine the effect of methods of instruction upon student achievement.

525 26) To distinguish between variables which effect memory and those which effect learning.

001 27) To state operational definitions for independent and dependent variables in a classroom learning environment.

354 28) To be able to recommend a program of instruction based upon the results of laboratory experimentation.

015 29) To be able to distinguish between a description and an explanation of educational achievement.

323 30) To propose a sampling scheme for testing a new curriculum on a statewide basis.

## Behavioral Objectives

At the level of behavioral objectives we need some standard way of representing conditions and behaviors as well as a means for determining how performance on tasks is to be judged. This latter step is accomplished by specifying a measure of performance for each set of behaviors and conditions. The following objectives are then written in a single format (one that is applied consistently throughout the proposal) which is designed to make the components of each objective as explicit as possible.

DAT Code 302

Conditions: Verbal reports of two experiments which test the effects of motivation on school performance.

Behavior: State the conditions under which the results of the two experiments may not be contradictory.

Measure: Evaluation by judges.

DAT Code 101

Conditions: A list of independent and dependent variables.

Behavior: Sort the list into two piles corresponding to independent and dependent variables respectively.

Measure: Speed of sorting and number of errors.

DAT Code 010

Conditions: A data table of scores relating an independent and dependent variable.

Behavior: Construct a graph to represent the relationship between the independent and dependent variable.

Measure: Evaluation by judges.

DAT Code 101

Conditions: Description of students learning American history.

Behavior: Write a list of potential dependent variables for assessing student learning.

Measure: Number of variables; evaluation by judges.

DAT Code 005

Conditions: The terms construct and theory.

Behavior: Write a description of how the two terms are related to one another in research work.

Measure: Evaluation by judges.

DAT Code 525

Conditions: Learning situation in which we wish to determine the effect on instructional methods upon ability to remember material.

Behavior: Construct a retention experiment which adequately controls for learning variables.

Measure: Evaluation by judges.

DAT Code 101

Conditions: Description of a system in which behavior is changed with time (e.g., students learning the capitols of states in a geography class).

Behavior: Write a list of potential independent variables for an experiment to study how the student learns.

Measure: Number of variables; evaluation by judges.

DAT Code 101

Conditions: A list of independent variables and a list of dependent variables.

Behavior: Select parts of items from the two lists that represent "scientifically reasonable" functional relationships.

Measure: Number of pairs; evaluation by judges.

DAT Code 315

Conditions: Statements containing descriptions of the nature of science.

Behavior: Recognize and label those which are true and those which are false.

Measure: Number of correct labelings.

DAT Code 101

Conditions: Statement of design to assess student achievement as a function of method of instruction.

Behavior: State appropriate controls.

Measure: Evaluation by judges.

DAT Code 101

Conditions: Statement of a school learning situation.

Behavior: State potential independent and dependent variables.

Measure: Evaluation by judges.

DAT Code 463

Conditions: Subject matter resources including a subject matter specialist.

Behavior: Produce a program of instruction which has a low error rate and high generalization to standard tests of achievement.

Measure: Evaluation by judges including subject matter specialists, psychologists, and students.

DAT Code 353

Conditions: A new curriculum.

Behavior: Produce a scheme for analyzing the results of testing a new curriculum against a standard curriculum.

Measure: Evaluation by judges.

DAT Code 323

Conditions: A new curriculum.

Behavior: Produce alternative sampling schemes for testing the new curriculum against a standard curricula on a statewide basis.

Measure: Number of schemes produced; evaluation by judges.

DAT Code 322

Conditions: Students learning geometry in a classroom setting.

Behavior: List possible extraneous variables for a study to determine the effectiveness of different methods of instruction.

Measure: Number of variables; evaluation by judges.

DAT Code 143

Conditions: A list of student scores on a test of subject matter achievement.

Behavior: Describe how these scores can be converted to facilitate interpretation.

Measure: Number of alternative conversions; evaluation of conversion by judges.

DAT Code 345

Conditions: Lists of student errors on two different programs of instruction to teach the same subject matter.

Behavior: Indicate the conditions under which each program would be most effective and why.

Measure: Evaluation by judges.

DAT Code 021

Conditions: Given the four levels of measurement: nominal, ordinal, interval and ratio.

Behavior: Produce examples of each kind of measurement.

Measure: Number of examples; evaluation by judges.

DAT Code 422

Conditions: A new method of counseling.

Behavior: Describe appropriate control groups for a study to determine the effectiveness of the new method.

Measure: Evaluation by judges.

DAT Code 422

Conditions: Description of an experiment to measure student learning as a function of motivational variables.

Behavior: List possible sources of confounding in the experiment and state the effect of these confoundings upon the interpretation of results.

Measure: Evaluation of judges.

DAT Code 040

Conditions: A data table of scores.

Behavior: Compute measures of central tendency and variability.

Measure: Number of measures computed; evaluation by judges.

DAT Code 050

Conditions: A data table of student scores on two tests of intelligence.

Behavior: Compute a coefficient of correlation between scores on the two tests.

Measure: Evaluation by judges.

DAT Code 465

Conditions: Description of the several studies with conflicting results to determine the effectiveness of instructional television.

Behavior: Interpret the results of the studies and make an educational recommendation.

Measure: Evaluation by judges.

DAT Code 465

Conditions: Results of several studies to determine effects of punishment as a motivational technique in instruction.

Behavior: Propose a theory of instruction based upon the results of the studies.

Measure: Evaluation by judges.

## Tasks for Training and Evaluation

This section contains tasks representing each kind of response: selection, recognition, and production. All of the tasks represent symbolic conditions (they should be thought of as written exercises rather than field problems).

DAT Code

### Multiple Choice (Selection)

005 1. One reason that the "scientific method" has received recent emphasis in the field of psychology is that (a) the common-sense or intuitive approach is usually incorrect, (b) one can make predictions with certainty when he correctly utilizes the scientific method, (c) recent philosophers have demonstrated that rationalism has no value in scientific pursuits, (d) the common-sense approach often leads to contradictory predictions.

005 2. As a science develops, (a) its theories become increasingly complex, (b) its theories become simplified as more phenomena become understood, (c) minute and detailed experimentation is replaced by more general and gross observation, (d) there is less emphasis on adequate explanation of the basic concepts.

005 3. The growth process of a science usually (a) is a slow, cumulative development, (b) combines the old methods of viewing the data with the new experimental techniques, (c) is an evolution from its past, resulting from discontent with the old, (d) is accompanied by an unchanging Zeitgeist.

003 4. If the accuracy of a prediction is demonstrated (a) there is normally only one explanation for the predictive relationship, (b) causation can then be inferred, (c) the predictive relationships are still open to alternative explanations, (d) the investigator can be assured that his study was both valid and reliable.

005 5. By a deterministic universe we mean (a) one in which man's efforts cannot change any part of the natural course of events, (b) one in which we predict from a probabilistic framework, (c) that if all factors influencing an event were known, then the event itself would be known with absolute certainty, (d) none of the above.

003 6. The dependability of an experiment in psychology (a) is its validity, (b) refers to the types of controls employed, (c) refers to the kinds of items being studied, (d) is its reliability.

300 7. Extraneous variables (a) are those which the experimenter manipulates, (b) are those which may be inadvertently brought into the study, (c) are those which the experimenter measures, (d) have effects only in the two-group experiment.

005 8. A probabilistic view of truth (a) allows us to predict with certainty, but only in given areas, (b) allows us to estimate the relative usefulness of various theories, (c) has recently received less emphasis in psychological science, (d) allows us to predict with only a degree of certainty.

001 9. "Intelligence is that which a properly standardized intelligence test measures" is an example of (a) a construct, (b) a nonsensical definition, (c) an operational definition, (d) an intuitive definition.

200 10. Randomization refers to (a) a method of control which is usually effective even with relatively small samples, (b) a method employed in developing theoretical constructs, (c) a type of generalization, (d) a method of control for which relatively large samples are needed.

005 11. Whitehead believed that (a) the emphasis of Western religion upon a rational divine being laid the groundwork for the acceptance of science, (b) Eastern science developed at a greater rate than did Western science, (c) accepting facts on the basis of religious faith has interfered with scientific development, (d) the effects of religion upon science were negligible.

403 12. We wish to test the effects of sleep deprivation upon psychomotor skills. Five Ss are tested immediately after waking; one week later they are retested following 24 hours sleep deprivation. Five other Ss are first tested after 24 hours sleep deprivation and retested the following week immediately after waking. This method of control is an example of (a) balancing, (b) counterbalancing, (c) equivalence, (d) randomization.

403 13. Probably the greatest opportunity for experiment bias to inadvertently creep into an experiment exists in (a) equivalence, (b) counterbalancing, (c) balancing, (d) statistical analysis of the data.

403 14. For which of the following is there the <u>least</u> amount of direct manipulation on the part of the experimenter? (a) Balancing. (b) Counterbalancing. (c) Equivalence. (d) Randomization.

005 15. According to Bridgman, explanation (a) is necessarily abstract and based upon converging evidence, (b) is the reduction of a situation to familiar elements which are accepted as a matter of course, (c) can never be complete, (d) must never be accepted as fact.

300 16. Counterbalancing refers to (a) an extraneous variable offsetting the effects of balancing in the experiment, (b) a method for increasing the reliability of a study, (c) a method to control for the effects of practice or fatigue, (d) a method of control usually restricted to the pilot study.

304 17. The wider the population from which an experimenter's samples are randomly drawn, the greater will be (a) the validity of the study and hence his ability to generalize, (b) the reliability of the study, (c) the number of subjects necessary to take part in his experiment, (d) the need to conduct a pilot study.

315 18. Extending the range of applicability, replication and theory testing are all (a) methods of experimental pursuit limited to the beginning investigator, (b) methods of scientific endeavor that give greater breadth to current knowledge, (c) methods that should be limited to the investigator of the original study, (d) types of pursuits that the "inspirational" experimenter should not concern himself with.

315 19. Psychological investigators attempt to use a wide range of experimental treatments (a) in studies extending the range of applicability of an earlier experiment, (b) usually only in studies using children or animals as subjects, (c) only in certain types of visual-discrimination studies, (d) in exploratory studies.

315 20. The maximizing of experimental effect is utilized in the hope that (a) it will enable one to generalize the findings to a greater population, (b) the wide range of experimental treatments will provide easily detectable differences, (c) it will increase the validity of the study, (d) it will decrease the need for an extensive review of the literature.

465 21. If the systematic approaches of scientific investigation were suddenly abandoned for the nonsystematic ones, what would be the most obvious loss to the scientific method? (a) Reliability. (b) Validity. (c) Its adaptability to changing times. (d) Its self-corrective nature.

465 22. If two studies attempting to investigate the same phenomena arrive at apparently contradictory findings, which question would we be <u>least</u> likely to ask? (a) Were the experiments in tune with the <u>Zeitgeist</u>? (b) What operational definitions were employed by the investigators? (c) What subject were used? (d) What were the experimental tasks?

412 23. A camp counselor wishes to conduct a two-group experiment. Half of the 40 youngsters attending summer camp are from a large metropolitan area and half are from a rural area. If the counselor used only one of these groups (10 Ss in each condition), he is employing the method of _____. If he assigns five metropolitan and five rural children to each condition he is using the method of _____. (a) equivalence, balance; (b) elimination, balance; (c) elimination, randomization; (d) equivalence, counterbalance.

415 24. In the example above, using both rural and metropolitan subjects will increase the _____, whereas using only one of these groups may increase the _____. (a) validity, reliability; (b) reliability, validity; (c) need for later replication, ability to generalize; (d) need for randomization, equivalence.

032 25. The frequency polygon is plotted from (a) the class limits, (b) the midpoints, (c) the number in each interval that has the greatest frequency, (d) this is arbitrary--it depends upon the preference of E.

032 26. We find that the X axis of our histogram extends 60 units when we plot our data, even though our actual range was only 55 and our interval width 5. This probably means that (a) we made an error when calculating our range, (b) a division error was made in determining the width of the intervals, (c) it is a natural phenomenon due to the extension for actual limits, (d) we should erase this final figure.

032 27. The boundary of each class interval is used for plotting the data in (a) a frequency polygon, (b) a histogram, (c) neither of the above, (d) both of the above.

032 28. We normally utilize the actual class limits in graphing psychological distributions instead of the working limits because (a) the former makes the categories have continguous boundaries, (b) we can determine the interval width by subtraction if we use the working limits, (c) both of the above, (d) neither of the above.

053 29. In psychology a significant difference is (a) one unlikely to have happened by chance alone, (b) one of practical import, (c) one that could not have happened by chance alone, (d) any difference large enough to be measurable.

032 30. In a bar graph depicting the populations of three states, the use of a nonzero origin would be most apt to (a) minimize the apparent differences, (b) maximize the apparent differences, (c) result in a shorter graph although apparent differences would remain approximately equal, (d) result in the appearance of a smaller population for all the states.

142 31. A positively skewed distribution is (a) one in which the majority of the scores are clustered at the lower end of the scale, (b) a normal curve of numerically low raw-score data, (c) one in which the majority of the scores are clustered to the right or higher end of the scale, (d) one skewed in the direction that had been hypothesized.

041 32. An administrator of a state hospital, wanting to present a picture of a high release rate of chronic schizophrenics, obtains a negatively skewed distribution when plotting the length of stay of 200 such admissions. If he wished to bias the finding in his favor, which measure of central tendency would he be most likely to report? (a) The mean. (b) The mode. (c) The median. (d) Either the mode or the median.

041 33. The standard deviation is (a) the square root of the sum of the deviations from $\bar{X}$ divided by N, (b) the square root of the variance, (c) the sum of the raw-score values divided by the total number of scores, (d) the square of the variance.

041 34. The sum of all deviations from the mean (a) varies in size according to the size of the original raw scores, (b) is equal to N for the particular sample, (c) is equal to zero, (d) is equal to X - M for the sample.

041 35. The "average deviation" (a) is occasionally utilized but is mathematically unsatisfactory, (b) is logically and mathematically unsound and is not used in the field of psychology, (c) is another name for the standard deviation, (d) is variance divided by N.

041 36. The mean squared deviation is usually referred to as (a) simply the mean, (b) the standard deviation, (c) the average deviation, (d) variance.

051 37. Of the mean and standard deviation, the two parameters of the normal curve, which of the following statements is true? (a) Distributions with identical means will have the same variance, and those with identical variance will have the same mean. (b) The two measures are independent of each other. (c) These parameters are insufficient, without additional information, to allow us to describe the shape of the distribution. (d) If two distributions have the same variance then their means will be equal, but the inverse of this relationhsip does not always hold true.

041 38. In a distribution with a mean of 50 and a $o$ of 6, if we wished to express the raw-score of 59 as a standard score (so many deviation units from the mean), we would do so in the following manner:

(a) $\frac{o}{X} = \frac{6}{-9} = -.67$ (b) $\frac{o}{x} = \frac{6}{9} = +.67$

(c) $\frac{x}{o} = \frac{9}{6} = +1.50$ (d) $\frac{x}{o} = \frac{-9}{6} = -1.50$

041 39. The formula $\frac{\Sigma/x/}{N}$ is a method of calculating the (a) variance, (b) standard deviation, (c) average deviation, (d) mean.

041 40. $\Sigma X^2$ indicates that we are to (a) sum all of the raw scores and square the total, (b) sum all of the squared deviations from the mean, (c) sum all of the deviations from the mean and square the total, (d) sum all of the squared raw scores.

142 41. We administer tests of reading speed to 200 entering freshman at Flunkout U. The scores obtained by these students form a normal distribution. Therefore, _____ people fall between $\frac{x}{o}$ scores of $\pm$ 1.96 and _____ fall between $\pm$ 2.58.
(a)99, 95; (b) 190, 198; (c) 95, 99; (d) 198, 190.

142 42. An achievement test is administered to six high-school English classes. The mean score is 80 and the $o$ is 14. Joe Doe receives a score of 101. This score is equal to a $\frac{x}{o}$ score of (a) + 1.5, (b) + 1.96, (c) -1.5, (d) -1.96.

142 43. If Joe Doe in the problem above had obtained a score of + 2.58 what percentage of the students would have scored above him? (a) 1%. (b) 5%. (c) 1/2 %. (d) 2.5%.

253 44. The mean score on a particular aptitude test for students at Flunkout U. is 80. We test 20 psychology students from this population and obtain a mean of 89 and a $o_{\bar{x}}$ of 3. (a) We could reject the null hypothesis. (b) We would have to test a second group of psychology students and compare the means of the two samples. (c) The null hypothesis would be confirmed. (d) The null hypothesis would not be applicable here because our sample was not a random representation of students at Flunkout U.

253 45. An investigator hypothesizes that children attending elementary school in a slum area of a large city will score lower on an IQ test than will the general population ($\bar{X}$ = 100, $o$ = 16). He tests 10 children and obtains a mean IQ score of 105. From these results (a) he may reject the null hypothesis and his hypothesis is confirmed, (b) his hypothesis is confirmed, (c) he may reject the null hypothesis, (d) he may neither reject the null hypothesis nor confirm his own hypothesis.

253 46. The mean annual income for high school students in town A is \$150 and the $\sigma$ is 25. We survey 25 students from this population and arrive at a sample mean of \$160. Our standard error of the mean for this sample is:

(a) $\frac{160 - 150}{\sqrt{25}} = 2$

(b) $\frac{25}{\sqrt{25}} = 5$

(c) $\frac{25}{160 - 150} = 2.5$

(d) $\frac{160 - 150}{25} = .4$

253 47. The null hypothesis states that (a) the mean of the sample and the $\sigma_{\bar{X}}$ of the population are the same, (b) any differences between the sample and population means are the result of chance, (c) the population from which the sample was drawn is not equal to the true population, (d) there are significant differences between the sample and population means.

253 48. An investigator sampled 50 individuals in a city in which the mean number of years of formal schooling for adults was 11. He found the mean for his sample to be 12 years schooling, and the $s_{\bar{X}}$ was .286. If he wished to determine the significance of his findings, which computations would he most likely employ?
(a) He would first have to calculate the $\sigma$ of the population,

(b) $\frac{\sigma_{\bar{X}}}{\sqrt{N}}$

(c) $\frac{12 - 11}{.286}$

(d) $\frac{\sigma_{\bar{X}}}{N - 1}$ and $\frac{\Sigma X^2 - (\Sigma X)^2/N}{N}$

253 49. An investigator hypothesizes that the mean IQ is higher than 100 among the general public living in college towns. He selects 400 individuals at random from eight college towns in his state and administers IQ tests. He obtains a mean IQ of 108. The steps in his analysis of this data would probably be: (a) computation of sample s, $s_{\bar{X}}$ using sample s, z test; (b) sample s, $\sigma_{\bar{X}}$ using population $\sigma$, t test; (c) sample s, $s_{\bar{X}}$ using sample s, t test; (d) sample s optional, $\sigma_{\bar{X}}$ using population $\sigma$, z test.

253 50. The mean number of hours college preparatory high school seniors spend in reading each month is 100 and the $\sigma_{\bar{x}}$ is 2. What score would be required to indicate a significantly greater amount of reading by any given student? (a) 103.92 hours, (b) 101.06 hours, (c) 196.00 hours, (d) 102.00 hours.

242 51. Two estimates of the variance of a single population should (a) fall within the range of ±2.58 $\sigma$ of that population, (b) differ according to the technique used to measure them, (c) be approximately equal, (d) differ in proportion to the homogeneity of the population from which they were sampled.

343 52. An investigator randomly assigns 10 college students each into three study groups (early morning, afternoon, late night) to determine if the period of the day at which people study has an effect on their retention. The Ss live in a controlled environment for one week, on the third day of which the experimental treatment (study of predetermined material) is administered. The seventh day the investigator tests for retention, and in computing his analysis he sees that his MS within groups is larger than his MS between groups. This indicates to him that (a) he has made an error in his calculations, (b) there was more than the expected variability between groups, (c) there was more variability between subjects within the same group than there was between groups, (d) that there should have been additional controls in his experiment.

343 53. The above investigator decides to reconduct his study with 30 elementary school children and compare the results. At retention he computes the average number of errors and finds: morning group, 5; afternoon, 2; night, 13. His F ratio is significant (16.0). He may conclude that (a) there is a significant difference between the three groups ($p < .01$) and $t = 4$, (b) there is a significant difference between all three groups ($p < .01$) and the afternoon group retains significantly more than does the night group ($p < .01$), (c) there is a significant difference in the scores of the afternoon and night groups ($p < .01$), (d) there is a significant difference in the individual scores of the three groups ($p < .01$).

445 54. Sloppiness in experimentation in a study to be analyzed by analysis of variance is most likely to result in (a) an increase in the within-groups variance, (b) a decrease in the F ratio, (c) a decrease in reliability, (d) all of the above.

445 55. If an investigator hired graduate students to serve as experimenters in his research without informing them as to his specific hypotheses, the most likely result would be (a) an increase in reliability, (b) a decrease in systematic extraneous variables, (c) a decrease in non-systematic extraneous variables, (d) a decrease in validity.

445 56. The scientific method (a) has some methodological limitations but few, if any, logical limitations, (b) is the only system to accurately order knowledge, (c) makes predictions from a probabilistic framework, (d) both "a" and "b" above.

445 57. In a true experimental study, the between-groups treatment difference is due to (a) the independent variable, (b) any differences in experimental room atmosphere that may exist between groups, (c) any difference in treatment of the groups by the experimenter that may exist, (d) all of the above.

445 58. A psychoanalyst is administering therapy to 20 adult males of approximately the same age, socioeconomic level, and intelligence. Eight of these men were raised in orphanages, and all eight of them tend to react to problems with anxiety and frustration. The other 12, raised by their own families, tend to react to problems with anger. The psychoanalyst concludes that being raised an orphan appears to cause a less aggressive reaction to problem or thwarting situations. (a) This is an ex post facto study for which causation could be inferred only from converging evidence. (b) This is an ex post facto study which could subsequently be conducted as a true experiment. (c) This is an experiment from which causation can be inferred. (d) This is a true experiment from which causation may be inferred only from converging evidence.

455 59. If an investigator conducts a study in which he successfully demonstrates that an increase in anxiety leads to a decrease in performance on a statistics test, his confirmed hypothesis becomes, for the population to which it can be generalized, (a) an operational definition, (b) an accepted fact, (c) a probabilistic statement, (d) a construct.

455 60. The systematic extraneous variable of experimenter bias is particularly difficult to control because (a) the person concerned will usually attempt to conceal the errors he makes in the experiment, (b) they most frequently simply result in nonsignificance, (c) no one has ever been able to adequately demonstrate their existence, (d) they are usually not detectable by standard types of data analysis.

455 61. The presence of extraneous variables in a study is most often reflected by (a) an increase in the within-groups variance, (b) an increase in the numerator of the F ratio, (c) a decrease in the between-groups variance, (d) a decrease in the within-groups variance.

455 62. When we employ a wide range of experimental treatments in an exploratory study, we (a) maximize the possible range of error variance, (b) maximize our experimental variance, (c) are attempting to minimize our experimental variance, (d) are increasing the possibility of a Type II error.

455 63. Error variance can be reduced through the use of (a) elimination, (b) extraneous variables, (c) theories, (d) none of the above.

455 64. We never "prove" the null hypothesis in research, because (a) a finding of no significant difference automatically proves it, (b) in psychology we are interested only in identifying areas of difference, (c) we are unable to place probability values on other factors which might have affected the results, (d) it is already proven and it is the experimentalist's task to disprove it.

465 65. An investigator administers three experimental treatments in a correlated-scores analysis of variance design. Eight Ss participate in these sessions. The degrees of freedom for this analysis would be (a) Treatment 3, Ss 8, error 24, total 35, (b) Treatment 2, Ss 8, error 11, total 21, (c) Treatment 2, Ss 7, error 9, total 18, (d) Treatment 2, Ss 7, error 14, total 23.

465 66. If eight different subjects had been used in each of the three treatment groups above (simple analysis of variance), the degrees of freedom would have been (a) between 2, within 7, total 23; (b) between 2, within 21, total 23; (c) between 3, within 7, total 10; (d) between 2, within 7, total 9.

465 67. If one employs matched rather than the same subjects in a correlated scores analysis of variance, (a) the Subjects MS would probably be smaller than it would be with the same subjects, (b) the subjects MS would probably be larger than it would be with the same subjects, (c) the error term would probably be smaller than when the same subjects were used, (d) the treatment effect would probably be larger than it would be with the same subjects.

465 68. The _____ in a simple analysis of variance is divided into the _____ in a correlated scores design. (a) Within SS, Subjects and error SS, (b) Total SS, Treatment and Subjects SS, (c) Between SS, Treatment and Subjects SS, (d) Within SS, Treatment and Subjects SS.

465 69. A correlated-scores analysis of variance is usually _____ than an analysis using independent scores, but the results of the latter analysis are _____ (a) more valid, more reliable, (b) less accurate, less valid, (c) more general, easier to interpret, (d) more accurate, more general.

465 70. In counterbalancing the order of treatment effects in a study employing three experimental conditions, if one uses all possible orders of these treatments, (a) both position and interaction effects are eliminated, (b) position effects, but not interaction effects are eliminated, (c) the number of Ss necessary to conduct the experiment doubles, (d) the data becomes too cumbersome to handle efficiently.

122 71. A standardized test of social adjustment is administered to 100 children attending elementary school in an impoverished area and 100 children from a middle-class school. The tests result in a numerical score for each child. Most probably, one could assume that scale of measurement for this data? (a) Nominal. (b) Ordinal. (c) Interval. (d) Ratio.

True-False (Recognition)

465 1. The validity of an experiment is unrelated to its reliability.

005 2. The early empiricists believed sensory experience to be the primary factor in the development of understanding.

005 3. Science may be regarded as our nearest approach to a universal language.

005 4. One of the limitations of science is its difficulty in adapting to changing times.

101 5. The dependent variable is that which the experimenter measures.

005 6. Causation can legitimately be inferred from prediction, but only when the accuracy of such prediction has been demonstrated.

005 7. Any complex logical system that we can construct in psychology may never be totally correct.

005 8. Psychologists try, whenever possible, to work from a framework of complete determinism.

200 9. For the randomization process to be effective, relatively large samples must be employed.

005 10. A constant interchange between observation and theorizing is necessary for the progress of science.

005 11. The nonsystematic approaches are more likely to be directly based upon the Zeitgeist than are the systematic approaches.

005 12. The primary criterion for judging the value of any research proposal is its pragmatic possibilities.

132 13. It is only necessary to use an origin of zero on the Y axis of a bar graph when this is the origin used for the X axis as well.

132 14. In research work, the independent variable of a study is usually placed on the Y axis of a graph.

132 15. Midpoints are determined by adding one-half the interval width to the lower working limit.

032 16. The range of a distribution is calculated by subtracting the lowest score of the data from the highest score possible for the particular measure.

032 17. One advantage of the frequency polygon is that the area under a given point is proportional to the number of Ss who have attained that particular score or lower.

041 18. Ninety-five percent of the cases forming a normal distribution fall within ±2.58 $\sigma$'s.

041 19. The standard deviation is the square root of the variance.

041 20. The distribution 1, 2, 2, 3, 4, 4, 5 is normally distributed.

043 21. An instructor assigns grades on the basis of the class curve. Performance is rather poor on the first exam, so all students in his class study harder for the next one. As a result, everyone obtains a score exactly 20 points higher than on the first test. The individual letter grades assigned will be the same on both exams.

043 22. On the third exam in the class mentioned above, all students except one obtain scores five points higher than on the second test. This individual obtains a perfect paper by achieving a score 20 points higher than on the second exam. His score will probably be the only one to change.

041 23. The median is the measure of central tendency most sensitive to change in a distribution.

041 24. $(\Sigma X)^2$ indicates that we are to sum all of the scores first and then square the sum.

242 25. The decrease in variability as we increase sample size is a linear one.

242 26. At any given time, the population mean should have only a small amount of variability.

242 27. If we obtain a z score value of 2.47 in computing the difference between a sample mean and the assumed population mean, we may say that the probability that the null hypothesis is true is less than .01.

122 28. The ratio scale can be assumed for the greatest number of present-day standardized psychological tests.

122 29. Most ordinal-scale data can also be expressed in terms of the interval scale but never in terms of the ratio scale.

262 30. Chi-square involves nonparametric ordinal-scale data.

## Completion (Production)

300 1. Psychologists have attempted to minimize the effects of not testing all possible variations in behavior through the use of _____ _____.

305 2. The use of control groups in an experiment often enables us to make some _____ relationship statement.

101 3. The _____ variable is that which the experimenter manipulates.

101 4. An ___ ___ ___ study is one in which the independent variable is not directly manipulated by the experimenter.

315 5. Use of both ascending and descending trials in the experiment on brightness discrimination in your text increased the _____ of the study.

302 6. The reading of standardized instructions to the subjects, random assignment of subjects to groups, and a set time element for completion of a task in an experiment are all examples of methods which _____ _____ _____ _____.

204 7. The _____, or _____ _____ _____ is increased by the use of both males and females in a study.

041 8. In the formula for computing the standard deviation, $\sigma = \frac{\Sigma X^2 - (\Sigma X)^2 / N}{N}$, the numerator is _____ _____ _____ _____ _____ _____.

041 9. In a skewed distribution, the measure of central tendency nearest the tail of the distribution is the _____.

041 10. A score of 64 in a distribution with a mean of 52 and a standard deviation of 8 would have a $\frac{x}{\sigma}$ score value of _____.

242 11. The standard error of the mean is simply the _____ _____ of the distribution of sample means.

242 12. _____ refers to all of the scores about which we wish to generalize, and _____ refers to a limited portion of this _____.

242 13. If the population of scores from which our sample originates is normally distributed, then _____ _____ and _____ ___ or _____ are the only two parameters influencing the standard error of the mean.

242 14. The standard deviation computed from a sample tends, on the average, to be _____ than the population $\sigma$ may be greatly _____.

122 15. In the _____ scale, the assignment of numbers is arbitrary.

## Short Answer (Production)

300 1. What is the primary effect of nonsystematic extraneous variables on a study?

003 2. What are three of the goals of psychological science?

465 3. What is meant by the validity of an experiment? By reliability?

302 4. How could it be possible for two experimenters to test the effects of motivation on school performance, obtain opposite results, and yet state that their studies were not, in fact, contradictory?

300 5. What is meant by (a) balancing, (b) counterbalancing?

312 6. Give at least three reasons why a review of the literature is important before conducting a psychological experiment.

312 7. What is meant by the statement that the use of multiple trials in a situation is one more example of the randomization process?

005 8. In what way does replication of studies currently in the literature contribute to the self-corrective nature of science?

242 9. What do we mean by "the variability of a sample mean"?

242 10. What is the null hypothesis and what does our rejecting it or not rejecting it imply?

242 11. How can we more closely approximate the population standard deviation from a sample s?

315 12. What is the logic behind the degrees-of-freedom concept, and how do we compute df for any given sample?

254 13. Which would we expect to be higher: the t value obtained through use of the sample s or the z value computed with the population $\sigma$? Why?

200 14. In 1936 the Literary Digest conducted a survey by randomly selecting names from telephone directories throughout the nation, and on the basis of this survey, announced that Landon would defeat Roosevelt in the presidential election. Roosevelt, however, defeated Landon by a sizable margin. From your knowledge of statistics, why do you suppose the results of the survey could not be successfully generalized to the entire population? That is, what requirement was unmet and why?

200 15. What would have been a more suitable technique for obtaining opinions in the survey above?

200 16. A large television corporation wishes to determine if a silent commercial is as effective in promoting memory of a product as is one in which the message is verbalized. One hundred names are randomly selected from the register of people who purchased color television sets within the past year, and these individuals are randomly divided into two equal experimental conditions. One group is shown the silent and one a regular commercial. A month after participation in this session, each S is given a list of 30 statements, 15 of which appeared in the commercial and 15 which did not. They are asked to check each bit of information that they remember, and a score is computed for each of them by subtracting the number incorrectly checked from the number correctly checked. Would conclusions drawn from this study be valid for all commercial television viewing? Why or why not?

## Essay (Production)

005 1. What is meant by (a) construct; (b) theory; and how are the two interrelated?

415 2. Describe what is meant by the statement that theory is both a tool and a goal. How do theories differ from simple summary statements, and what do you think is the significance of the interchange between observation and theorizing?

254 3. Differentiate between uses of, and assumptions for, the z test and the t test.

463 4. If you recorded the scores received on an IQ test administered to 100 individuals at a state college and found the mean to be 115 on this test, you could best determine if this differed significantly from the population mean through the use of what statistic?

463 5. Ten rats, previously trained to jump, are tested in a Lashley Jumping Apparatus and latency of jump recorded as a function of either 24- or 48-hour food deprivation. All animals are tested under each level of deprivation. How could we determine if the relative performance of Ss changes in the two conditions.

463 6. The above investigators replicate their study, this time measuring the force or jump as a function of the 24- or 48-hour deprivation (use of a pendulum attached to the jumping platform yields a kymograph recording of the magnitude). How might they determine if there was a significant difference in performance in the two deprivation conditions?

041 7. Raw scores from standardized tests are frequently converted to percentile ranks, grade equivalents, or standard scores for purposes of interpretation. Explain how each type of conversion is performed and the circumstances under which one system would be preferable to another for a particular test. Cite examples of each.

051 8. Give three examples of an incorrect interpretation of the Pearson product-moment correlation coefficient. Explain what is wrong with the interpretation offered in those situations.

051 9. The product-moment correlation coefficient (r) is frequently employed simply as a descriptive index of relationships in a sample of paired observations. Give some interpretations of the value of an obtained r. What does it mean to state that an r of .25 is statistically significant at the .05 level?

051 10. Distinguish between the correlation and regression approaches to a problem. State the advantages and disadvantages, similarities and differences of the two approaches. Contrast two situations, one of which would call for the correlation approach, and one which would call for the regression approach.

154 11. Define the four levels of measurement used by Stevens (nominal, ordinal, interval, and ratio), give the formal properties of scales at each level, indicate the operations permissable in each one, and cite some statistics which might be considered appropriate with each type of scale.

255 12. Discuss random sampling with and without replacement and relate it to the concept, conditional probability. In the usual inferential situation in which the population size is a very large but finate number, which method is actually called for? Is this approach followed?

365 13. In the analysis of frequency data, review the various methods for analyzing data summarized in a 2x2 table. Explain the cautions, assumptions, and remarks that seem appropriate for a person to consider if he were about to use this approach.

355 14. One technique frequently employed in making statistical inferences is the estimation of population parameters from sample statistics. Several properties of estimators are often discussed. These include unbiasedness, efficiency, consistency, sufficiency. Choose two of these to discuss. Why or in what sense is each a desirable property? Give an example of a statistic which possesses the properties you have discussed.

165 15. Distinguish as clearly as possible between the class of statistical procedures known as parametric methods and the class of techniques known as nonparametric methods. In your discussion indicate the relevance of assumptions regarding both measurement and statistical models. Give an example of a parametric procedure and a nonparametric analog.

305 16. Discuss the function of control groups or comparison groups in an experimental study, how such groups might be constituted, and how one might determine their adequacy in a particular study.

405 17. The problem of confounding enters into the interpretation of results from the Latin square design in a special way. Explain what is meant by confounding in general and how it enters into the Latin square design specifically. Provide an example.

465 18. Discuss the nature and function of analysis of covariance. Contrast that technique with treatments x levels analysis of variance in terms of circumstances in which each is applicable or preferable, nature of answer provided, and theoretical assumptions involved.

465 19. Some people have advocated that testing of hypothesis be abandoned and that results of research be reported in the form of confidence intervals rather than statements of significance. Consider the pro's and con's of such a suggestion.

165 20. Three somewhat different approaches to the development of a theory of human behavior may be seen in the techniques of hypothesis testing, parameter estimation, and factor analysis. Explain briefly how each technique is used and how it can contribute to the development of theory.

165 21. What are the assumptions underlying the usual multiple regression model when tests of significance at various stages of the analysis are to be used? Give an example of a situation in which such a procedure might be used. What is cross-validation? How would it relate to the example given?

423 22. Outline the general considerations which you would consider important in developing a large scale investigation of the effects of certain types of pre-trial publicity on the outcome of jury trials. Consider some or all of the following factors:

1. News media (TV, radio, newspaper)
2. Type of item (arrest and charge, confession, conduct during arrest, etc.)
3. Selection of panels of jurors from community to participate in simulated trials.
4. Video tape vs. live drama for juror's examination.
5. What would be appropriate criterion measure, e.g., guilt or innocence as established by jury panel.
6. Number of such juries which would be required.

Present some design (repeated-measures?) which would contain those elements which you consider most important to be controlled.